MCMC sampling colourings and independent sets of $G(n, d/n)$ near uniqueness threshold

Charilaos Efthymiou*
Goethe University, Mathematics Institute, Frankfurt 60054, Germarny
efthymiou@math.uni-frankfurt.de

Abstract

Sampling from *Gibbs distribution* is a central problem in computer science as well as in statistical physics. In this work we focus on the k*-colouring model* and the *hard-core model* with fugacity λ when the underlying graph is an instance of Erdős-Rényi random graph $G(n, p)$, where $p = d/n$ and d is fixed.

We use the *Markov Chain Monte Carlo* method for sampling from the aforementioned distributions. In particular, we consider *Glauber (block) dynamics*. We show a dramatic improvement on the bounds for *rapid mixing* in terms of the number of colours and the fugacity for the corresponding models. For both models the bounds we get are only within small constant factors from the conjectured ones by the statistical physicists.

We use *Path Coupling* to show rapid mixing. For k and λ in the range of our interest the technical challenge is to cope with the high degree vertices, i.e. vertices of degree much larger than the expected degree d. The usual approach to this problem is to consider block updates rather than single vertex updates for the Markov chain. Taking appropriately defined blocks the effect of high degree vertices somehow diminishes. However, devising such a construction of blocks is a *highly non trivial* task.

We develop for a first time a weighting schema for the paths of the underlying graph. Vertices which belong to "light" paths, only, can be placed at the boundaries of the blocks. Then the tree-like local structure of $G(n, d/n)$ allows the construction of simple structured blocks.

Interestingly enough, the weighting schema captures so well the desired properties of the blocks that we do not even need to argue explicitly on *spatial mixing of Gibbs distribution* for the range of k, λ we consider. However, we do believe that our approach has further consequences as far as correlation decay is regarded. Thus, apart from the standard path coupling analysis, the main bulk of the analysis we use is of probabilistic nature.

*Partially supported by EPSRC grant EP/G039070/2.

1 Introduction

Sampling from *Gibbs distribution* is a central problem in computer science as well as in statistical physics. Examples include sampling from the uniform (or a weighted) distribution over combinatorial structures like k-colourings, independent sets, matchings of a given graph G. In this work we focus on colourings and independent sets when the underlying graph is an instance of Erdős-Rényi random graph $G(n,p)$, where $p = d/n$ and d is 'large' but remains bounded as $n \to \infty$. We say that an event occurs *with high probability (w.h.p.)* if the probability of the event to occur tends to 1 as $n \to \infty$.

For this kind of problems the most powerful algorithms and somehow the most natural ones are based on the Markov Chain Monte Carlo (MCMC) method. The setup is an ergodic, time-reversible Markov chain over the k-colourings (or independent sets) of the underlying graph. The updates guarantee that the equilibrium distribution of the chain is the desired one. Here we use standard *Glauber block updates* in the course of this paper we refer to the chains as *Glauber dynamics*. The main technical challenge is to establish that the underlying Markov chain has *rapid mixing*, i.e. it converges sufficiently fast to the equilibrium distribution (see [9, 16, 15]).

Given the input graph $G(n, d/n)$, the focus is on two distributions. The first one is the *colouring model*, i.e. the uniform distribution over the k-colourings of the input graph. The second one is the *hard-core model* with *fugacity* λ, i.e. the independent set σ is selected with probability proportional to $\lambda^{|\sigma|}$. The parameters of interest are the number of colours k and the fugacity λ. The aim is to show rapid mixing for k as small as possible and λ as large as possible for the corresponding models.

For MCMC algorithms to converge, typically, the bounds for both k and λ are expressed in terms of the *maximum degree* of the underlying graph. Examples of such bounds are [5, 10, 11, 13, 20, 21, 24] for colouring and [6, 7, 20, 25] for independent sets. In that terms, what makes the case of $G(n, d/n)$ special is the (relatively) big fluctuations in the degree of the vertices. To be more specific, w.h.p. the vast majority of vertices in $G(n, d/n)$ are of degree close to d, while the maximum degree is as huge as $\Theta\left(\frac{\ln n}{\ln \ln n}\right)$. In such a situation, it is natural to expect that the rapid mixing bounds for both k, λ depend on the *expected degree* d, rather than maximum degree.

Sophisticated but mathematically non rigorous arguments from statistical physics (e.g. in [18]) support this picture. They suggest that w.h.p. over the instances of $G(n, d/n)$ the Glauber (block) dynamics on k-colouring has rapid mixing for any $k > d$. Furthermore, for $k < d$ the chain is expected to be non-ergodic and only weaker notions of convergence hold. To our knowledge, there are no predictions for the fugacity as far as the hard-core model is concerned. However, using the result in [17] and standard arguments we could conjecture that we have rapid mixing as long as $\lambda < \frac{(d-1)^{d-1}}{(d-2)^d} \approx \frac{e}{d}$.

The best bounds for Monte Carlo sampling appear in [22] (which improved on [4]). The authors in [22] provide for a first time rapid mixing bounds for k and λ which depend on the expected degree d. That is, w.h.p. over $G(n, d/n)$ there are functions $f(d)$ and $h(d)$ such that Glauber dynamics has rapid mixing for k-colourings and hard-core for $k \geq f(d)$ and $\lambda \leq h(d)$, respectively[1]. However, the values for k and λ that are allowed there are many orders of magnitude off the conjectured bounds. Here we improve on these bounds dramatically. We show that w.h.p. over the underlying graph $G(n, d/n)$ we have rapid mixing for $k \geq \frac{11}{2} d$ and for $\lambda \leq \frac{1-\epsilon}{2}\frac{1}{d}$. That is, we approach the conjectured bounds for rapid mixing only within small constants.

Remark. The main strength of MCMC approximate sampling algorithms relies on the fact that the configuration space should be, somehow, "well-connected". Rapid mixing allows approximating the target Gibbs distribution within distance err by letting the algorithm run for $\ln(err^{-1}) \times poly(n)$ steps. There can be approximate sampling algorithms based on weaker assumptions. In [8] the author of this work proposed such an algorithm for approximate sampling k-colourings of $G(n, d/n)$ which has a

[1] Even though these functions are not given explicitly it is conceivable from the analysis that it holds that $f(d) \geq d^c$ and $h(d) \leq d^{-c'}$ for fixed $c, c' > 0$.

notable performance in terms of minimum k, as it requires $k \geq (2+\epsilon)d$. This algorithm is not a Monte Carlo one. Of course the approximation guarantees there are weaker. That is, the error of the output is a vanishing function of the size of the graph n, i.e. for a fixed sized graph the error is fixed. On the other hand, for the algorithms here the error can be *arbitrarily* small by allowing the algorithm to run for a sufficiently large number of steps. Furthermore, the approach we follow here applies to models other than colouring, e.g. hard-core model.

We use the well-known Path Coupling technique, from [3], to show rapid mixing. Path Coupling is also used in both of the previous papers on the problem, i.e. [4, 22]. For k and λ in the range of our interest the technical challenge is to cope with the high degree vertices, i.e. vertices of degree much larger than d. The natural approach is to consider block updates rather than single vertex updates for the Markov chain. In particular we use the observation that the effect of high degree vertices somehow diminishes when they are away from the boundary of their block. Devising such a block construction is a *highly complex* task. We introduce for a first time a weighting schema for the paths of the underlying graph which allows a desired block construction.

To be more concrete, we use our weighting schema to assign weight to each path in $G(n, d/n)$. These weights allow to distinguish which vertices can be used for the boundaries of the blocks. We call such vertices as *break-points*. A break point should have all the paths emanating from it of sufficiently small weight. It turns out that w.h.p. there is a plethora of break-points in $G(n, d/n)$. This allows creating small, simple structured blocks. For further discussion on the weighting schema see Section 1.3.

Remark. In the full version of this work we study some extra models which are of main interest in statistical physics, e.g the Ising model and the Potts model with positive temperature.

Notation We use small letters of the greek alphabet to indicate colourings or independent sets, e.g. σ, τ. Also, by $\sigma(v)$ we indicate the assignment of the vertex v under the configuration σ. For a vertex set B we call *(outer) boundary* of B the vertices outside B which are adjacent to some vertex inside B.

1.1 The Models and the Algorithm

We consider the *colouring model* and the *hard core model*. For each of these two models we consider a graph $G = (V, E)$ and a set of *spins* C. We define a *configuration space* $\Omega \subseteq C^V$. Given Ω, the model specifies a distribution $\mu : \Omega \to [0, 1]$. This distribution is usually called *Gibbs distribution*.

Colouring Model. Given a graph $G = (V, E)$ and a sufficiently large integer k, the *colouring model* specifies the following: The configuration space Ω is the proper k-colourings of G. The Gibbs distribution is the uniform distribution over Ω. That is, for $\sigma \in \Omega$ it holds that

$$\mu(\sigma) = \frac{1}{|\Omega|}.$$

Hard Core Model. Given $G = (V, E)$ and a real $\lambda > 0$, the hard core model with *fugacity* λ is as follows: The configuration space Ω is all the independent sets of G. The Gibbs distribution specifies that for each $\sigma \in \Omega$, $\mu(\sigma)$ is proportional to $\lambda^{|\sigma|}$, where $|\sigma|$ is the cardinality of σ. That is

$$\mu(\sigma) = \frac{1}{Z}\lambda^{|\sigma|},$$

where Z is a normalizing quantity, i.e. $Z = \sum_{\sigma \in \Omega} \lambda^{|\sigma|}$. Z is usually called *partition function*.

In this work we propose a Markov Chain Monte Carlo algorithm for sampling from the two models above. The general form of the algorithm is as follows:

Algorithm: The input is a graph $G = (V, E)$, a number $err > 0$ which is the error in the distribution of the output sample and a set of parameters [2] which specify the target Gibbs distribution μ.

First, the algorithm partitions the set of vertices V into an appropriate set of blocks $\mathcal{B}$. Given $\mathcal{B}$, it simulates the following Markov chain and returns the configuration of the chain after $T = T(err)$ steps.

- Start from an *arbitrary* configuration.
- At each transition, the chain chooses uniformly at random (u.a.r.) a block $B \in \mathcal{B}$. If X_t is the current state of the chain, then the next one, X_{t+1}, is acquired as follows:
  - For every vertex $u \notin B$ set $X_{t+1}(u) = X_t(u)$.
  - Set $X_{t+1}(B)$ according to Gibbs distribution conditional that the spins outside B are set $X_{t+1}(V \backslash B)$.

Remark. Given a set of technical conditions known as *ergodicity*, it is trivial to show that the above chain converges to the appropriate Gibbs distribution μ.

To deal with the time efficiency and the accuracy of our algorithm we need to address the following issues:

1. The construction of $\mathcal{B}$ is done in polynomial time.
2. There is an algorithm that gives in polynomial time an initial configuration of the chain.
3. Each transition of the chain can be implemented in polynomial time.
4. The chain converges to stationarity sufficiently fast, i.e. we have rapid mixing.

We use *mixing time*, τ_{mix}, as a measure of the speed of convergence of Markov chains. The mixing time is defined as the number of transitions needed in order to guarantee that the chain starting from an arbitrary configuration, is within total variation distance $1/e$ from the stationary distribution (see [19]).

Remark. In our context, we say that a Markov chain is *rapidly mixing* if τ_{mix} is polynomial in n.

Remark. It is not hard to see that the number of transitions that are required to get within error err from the stationary distribution is $T(err) \geq \ln\left(\frac{1}{err}\right) \times \tau_{mix}$.

1.2 Our Results

The main result of this work is in the following theorem.

Theorem 1 *Let d be sufficiently large and let $\epsilon > 0$ be fixed. On input $G(n, d/n)$ and sufficiently small $err > 0$ for the algorithm in Section 1.1 the following holds:*

colouring model: *W.h.p. over the input graphs $G(n, d/n)$ and for any $k \geq \frac{11}{2}d$ the algorithm returns a k-colouring of the input graph distributed within total variation distance err from Gibbs distribution. The time complexity scales as $\ln\left(\frac{1}{err}\right) \times O(n^c)$, for a fixed $c > 0$.*

hard-core model: *W.h.p. over the input graphs $G(n, d/n)$ and for any $\lambda \leq \frac{1-\epsilon}{2d}$ the algorithm returns an independent set of the input graph distributed within total variation distance err from Gibbs distribution. The time complexity scales as $\ln\left(\frac{1}{err}\right) \times O(n^{c'})$, for a fixed $c' > 0$.*

[2] e.g. number of colours k for colouring, fugacity for hard core

Theorem 1 follows as a corollary from a sequence of results that we present in this section. First, we need to remark the following: Up to this point we have not provided any details about the set of blocks $\mathcal{B}$ the Markov chain uses. These details are a bit technical. However, what is important, for the moment, is that each of the blocks in $\mathcal{B}$ should satisfy certain conditions which depend on the structure of the underlying graph G. In the following two theorems, the convergence holds w.h.p. over the instances of $G(n, d/n)$ because (among other restrictions) the underlying graph should admit the appropriate partition $\mathcal{B}$

Theorem 2 *Let $\mathcal{M}_c$ be the Markov chain on the k-colourings of $G(n, d/n)$ we define in Section 1.1. With probability $1 - o(1)$ over the graph instances $G(n, d/n)$ and for $k \geq \frac{11}{2}d$, the mixing time of $\mathcal{M}_c$ is $O(n \ln n)$.*

Theorem 3 *Let $\mathcal{M}_{hc}$ be the Markov chain of the hard core model on $G(n, d/n)$ with fugacity λ, we define in Section 1.1. Let $\epsilon > 0$ be fixed. With probability $1 - o(1)$ over the graph instances $G(n, d/n)$ and for $\lambda \leq \frac{1-\epsilon}{2d}$ the mixing time of $\mathcal{M}_{hc}$ is $O(n \ln n)$.*

The following result, Theorem 4, shows that w.h.p. we can construct the set $\mathcal{B}$ efficiently. Also, it shows that w.h.p. the structure of each block in B is quite simple. The proof of Theorem 4 appears in Section 2, once we have described in full detail how do we get the set $\mathcal{B}$.

Theorem 4 *With probability $1 - o(1)$ over the instances of $G(n, d/n)$ the following is true: There is a small fixed $s > 0$ such that the block construction can be made in time $O(n^s)$. Each block in $\mathcal{B}$ is either a tree or a unicyclic graph.*

Theorem 4 implies that, w.h.p. the set $\mathcal{B}$ is such that the updates in both chains ($\mathcal{M}_c$ and $\mathcal{M}_{hc}$) can be implemented efficiently by using dynamic programming (D.P.). The method is standard and it is based on reducing the update to computing the so called *partition function* for a block B with fixed boundary. That is we use D.P. to compute the partition function for each block B. The extra edge in the unicyclic blocks should not pose any particular problem in the D.P.[3].

Finally, for the initial state of the chain we work as follows: For $\mathcal{M}_{hc}$ we can trivially consider the empty independent set as the initial state. For the chain $\mathcal{M}_c$ we can get an initial state by using the algorithm suggested in [12]. The authors there provide a greedy, polynomial time algorithm which k-colours typical instances of $G(n, d/n)$ for any $k \geq (1 + c) \ln d/d$ and any fixed $c > 0$.

1.3 Convergence - Proof technique

We show rapid mixing by using the well-known *Path Coupling* technique in [3]. The technique goes as follows: W.l.o.g. we consider the colouring model. Assume that the underlying graph $G = (V, E)$ is of maximum degree Δ and, for the moment, let $k > \Delta$. Finally, assume that we have the Markov chain M on the k-colourings of G with single vertex updates.

Consider any two copies of M at state X_0, Y_0, respectively. We take X_0, Y_0 so that they have exactly one disagreement, i.e. their Hamming distance $H(X_0, Y_0)$ is equal to 1. The coupling carries out one transition of each copy of M. Let X_1, Y_1 be the colouring after each transition, respectively. A sufficient condition for rapid mixing is to exist a coupling for the transitions of the two copies of M such that

$$E[H(X_1, Y_1)|X_0, Y_0] \leq 1 - \Theta\left(n^{-1}\right). \tag{1}$$

To study the technique further, assume now that for $w \in V$ we have $X_0(w) \neq Y(w)$. It is natural to use a coupling that updates the same vertex in both copies. The cases that matter are only those where the coupling chooses to update either the disagreeing vertex w or one of its neighbours. If the update

[3] For a more detailed treatment on how the update should be, the reader is referred to [4].

involves the vertex w, then we get that $X_1 = Y_1$. This happens with probability $1/n$, where $|V| = n$. On the other hand, if the update involves a neighbour of w, then X_1, Y_1 may have an extra disagreement. In particular, the update of a neighbour of w can generate an extra disagreement with probability at most $\frac{1}{k-\Delta}$. Since the disagreeing vertex w has at most Δ neighbours, the probability of having an extra disagreement is at most $\frac{\Delta}{n}\frac{1}{k-\Delta}$. For $k \geq 2\Delta + 1$, it is direct that (1) is satisfied.

W.h.p. $G(n, d/n)$ is of maximum degree $\Theta\left(\frac{\ln n}{\ln \ln n}\right)$. That is, a vanilla path coupling would require an unbounded number of colours. Otherwise, i.e. if k is smaller than the maximum degree, there is no control on the expected number of disagreements generated in the coupling. However, it is possible to gain some control on the expected number of disagreements by using (appropriate) block updates rather than single vertex updates in the Markov chains. In particular, the blocks should be constructed in such a manner that the high degree vertices are somehow "hidden" inside the blocks [4].

In our setting, we consider two copies of $\mathcal{M}_c$ at states X_0, Y_0. The states differ only on the assignment of the vertex w. The coupling chooses uniformly at random a block B from the set of blocks $\mathcal{B}$ and updates the colouring of B in both chains. It turns out that the crucial case for proving (1) is when the outer boundary of B is not the same for both chains, i.e. the vertex w is at the outer boundary of the block B [5]. There, we need to upper bound the expected number of *disagreements*, the vertices which take different colour assignments after the update of colouring. The construction of $\mathcal{B}$ should minimize the expected number of disagreements.

To this end, we improve on an idea from [4]. We use the well-known "*disagreement percolation*" coupling construction, [2]. The disagreement at the boundary prohibits identical coupling of $X_1(B)$ and $Y_1(B)$. The disagreement percolation assembles the coupling in a stepwise fashion moving away from w. Disagreements propagates into B along paths from w. A disagreement at vertex $u' \in B$ at distance r from w propagates to a neighbour u at distance $r+1$ if $X_1(u) \neq Y_1(u)$. The disagreement percolation is dominated by an independent process such that each vertex $v \in B$ is disagreeing with probability

$$\varrho_v = \begin{cases} \frac{2}{k-(1+\alpha)d} & \text{if } \Delta(v) \leq (1+\alpha)d \\ 1 & \text{otherwise,} \end{cases}$$

where $\alpha > 0$ is some small constant and $k \geq 5.5d$. The disagreement propagates over the path L that start from w with probability at most $\prod_{u \in L \setminus \{w\}} \varrho_u$. The expected number of disagreements is at most the expected number of paths of disagreements that start from w and propagate inside B.

Intuitively, high degree vertices are expected to have an increased contribution to the number of disagreements. Mainly this is due to the following reason: If a high degree vertex is disagreeing, it has an increased number of neighbours to propagate the disagreement. However, for typical $G(n, d/n)$ and $k \geq \frac{11}{2}d$, it turns out that the larger the distance between a high degree vertex from w the less probable is for the disagreement to reach it. This, somehow, can balance the increased contribution that high degree vertices have. We exploit this observation in the block construction so as to control the overall number of disagreeing vertices in the updates.

To be more concrete, we introduce a weighting schema as follows: Each vertex u, of degree $\Delta(u)$ in $G(n, d/n)$, is assigned weight $W(u)$ such that

$$W(u) = \begin{cases} (1+\gamma)^{-1} & \text{if } \Delta(u) \leq (1+\alpha)d \\ d^c \cdot \Delta(u) & \text{otherwise,} \end{cases} \tag{2}$$

for appropriate real numbers $\alpha, \gamma, c > 0$. Given the weights of the vertices, each block $B \in \mathcal{B}$ should satisfy the following two properties:

(a) B is either a tree or a unicyclic graph

[4] This is the approach that is used in the analysis in both [4, 22].

[5] $w \notin B$ but w is adjacent to vertices in B

(b) For every path L between a vertex at the outer boundary of B and a high degree vertex inside B it should hold that $\prod_{u \in L} W(u) \leq 1$.

The property (a), above, guarantees that the updates can be implemented efficiently. The property (b) guarantees that the disagreement percolation coupling construction we describe above satisfies (1).

In the weighting schema, observe that the low degree vertices reduce the weight of the path L, while the high degree vertices increase it. Restricting the weight of a path between a high degree vertex in the block B and a boundary vertex, somehow, guarantees that the high degree vertices are sufficiently far from the boundary. I.e. so as to keep the weight of the path low we require a sufficiently large number of low degree vertices between the boundary vertex and the high degree vertex.

Remark. Taking a path L in $G(n, d/n)$ we expect that the product of the weights of its vertices is rather low. Mainly, this is due to the fact that only a very small fraction of vertices have large weight. That is, for each $u \in L$ it holds that $Pr[\Delta(u) > (1+\alpha)d] \leq \exp\left(-\alpha^2 d/3\right)$.

For a path L in $G(n, d/n)$ with $|L|$ vertices, we show that the upper tail of its weight is heavy. I.e. Theorem 8 implies that the probability that the weight of L is greater than 1 is at most $\exp(-d^{0.8}|L|)$.

Remark. We follow exactly the same approach to show rapid mixing of Markov chains for the hard core model. The partitioning of the vertices of the underlying graph is exactly the same. The only difference is the probabilities that the disagreement propagate inside the block B.

2 Block Creation & Proof of Theorem 4

The process of creating the set of blocks $\mathcal{B}$ considers three positive real numbers α, γ and c as parameters, i.e. $\mathcal{B} = \mathcal{B}(\alpha, \gamma, c)$. Consider α, γ and c to be fixed numbers, i.e. independent of d. Their exact value will be specified later. Consider, also, the graph $G(n, d/n)$ and the weighting scheme in (2) for its vertices. Given the weights, we introduce the concept of "influence".

Definition 1 (Influence) *For a vertex v, let $\mathcal{P}(v)$ denote the set of all paths of length at most $\frac{\ln n}{d^{2/5}}$ that start from v. We call "influence" on the vertex v, denoted as $E(v)$, the following quantity:*

$$E(v) = \max_{L \in \mathcal{P}(v)} \left\{ \prod_{v \in L} W(v) \right\}.$$

A vertex v is considered to be under no influence if $E(v) \leq 1$. The vertices under no influence are special for the blocks construction as they are used to specify the boundaries of the blocks.

Definition 2 (Break-Points & Influence Paths) *A vertex v such that $E(v) \leq 1$ is called* "break-point". *Also, a path L that does not contain break-points is called* "influence path".

Block Creation: We have two different kinds of blocks. For this, let $\mathcal{C}$ denote the set of all cycles of length at most $4\frac{\ln n}{\ln^5 d}$ in $G(n, d/n)$.

1. For each $C \in \mathcal{C}$ we have a block which contains every vertex $v \in C$ as well as all the vertices that are reachable from v through an influence path that does not use vertices of $C \backslash \{v\}$.

2. The remaining blocks are created as follows: Pick a vertex v which does not belong to any block, so far. Consider as a new block the vertex v and all the vertices that are reachable from v through an influence path.

Remark. From the construction it is apparent that each break-point is a single vertex block.

The reader may have noticed that there is no "heavy" path of length at most $\frac{\ln n}{d^{2/5}}$ that starts from a break point. However, in Section 1.3 we mention that the vertices at the boundaries of each block B should not be connected to a high degree vertex inside B with a heavy path. That is from the above definition of break-points and the block construction there may be undesirable heavy paths between the boundary of a block B and some of its vertices. It turns out that, w.h.p. none of the blocks in $\mathcal{B}$ has such a path.

2.1 Proof of Theorem 4

For typical instances of $G(n, d/n)$, it turns out that the influence paths we need to consider for the creation of the set of blocks $\mathcal{B}$ are rather short. Consequently, each of the blocks that are created is either a tree or a unicyclic graph. To be more concrete, the situation is as follows: W.h.p. cycles as those in the set $\mathcal{C}$ are far apart from each other in $G(n, d/n)$, e.g. at distance greater than $\ln n/\ln^2 d$. We are going to show that all the influence paths we need to consider are of much shorter length, e.g. at most $\ln n/\ln^5 d$. Therefore, no two cycles will be connected and no new cycles will emerge.

In the statement of the following theorem, we call *elementary* every path L in $G(n, d/n)$ such that there is no cycle shorter than $10 \ln n/d^{2/5}$ which contains two vertices of L.

Theorem 5 *Consider $G(n, d/n)$ and the fixed numbers $\gamma, c > 0$, $\alpha \in (0, 3/2)$. Let the set $\mathbb{U}$ contain all the* elementary *paths in $G(n, d/n)$ of length $\frac{\ln n}{\ln^5 d}$ that do not have any break-point. It holds that*

$$Pr[\mathbb{U} \neq \emptyset] \leq n^{-\frac{1}{3}\frac{\gamma}{1+\gamma}\ln d}.$$

The proof of Theorem 5 appears in Section 7.

Remark. From Theorem 5 and our previous remark about the mutual distance of the cycles in $\mathcal{C}$ the following observation is direct: W.h.p. the construction of $\mathcal{B}$ considers only influence paths that are *elementary*.

Using Theorem 5 it is direct to prove the following two lemmas.

Lemma 1 *With probability at least $1-2n^{-3/4}$ over the instances $G(n, d/n)$, the blocks that are created are either trees or unicyclic graphs.*

The proof of Lemma 1 appears in Section 13.1. Also, we have the following lemma holds.

Lemma 2 *With probability at least $1 - n^{-3/4}$ over the instances of $G(n, d/n)$, the following is true: There is a small fixed $s > 0$ such that the block construction can be made in time $O(n^s)$.*

The proof of Lemma 2 appears in Section 13.3.

Remark. Theorem 4 follows as a corollary of the two lemmas above.

3 Convergence

3.1 Colouring Model - Proof of Theorem 2

Ergodicity. For Theorem 2 first we need to consider when $\mathcal{M}_c$ is *ergodic*. From [4] we have that the Glauber dynamics (and hence the block dynamics we have here) is ergodic probability $1 - o(1)$ over the instances $G(n, d/n)$ when $k \geq d + 2$.

Remark. The proof for ergodicity in [4] goes as follows: It is shown that if a graph G has no t-core[6], then for all $k \geq t+2$ the Glauber dynamics for k-colouring yields an ergodic Markov chain. Then the authors use the result in [23], which states that w.h.p. $G(n, d/n)$ has no t-core for $t \geq d$.

For our underlying graph $G(n, d/n)$ we require it to have, additionally, the following three properties: (A) It can be coloured with d colours or less. (B) Let $\mathcal{B} = \mathcal{B}(\alpha, \gamma, c)$ be such that $0 < \gamma \leq \alpha \leq 10^{-2}$ and $c = 10$. Then, each of the blocks in $\mathcal{B}$ is either a tree or a unicyclic graph. (C) For the values of α, γ, c as in (B) the following holds: For each $B \in \mathcal{B}$ and every u at the outer boundary of B there is no path L from u to a high degree vertex $u' \in B$ such that $\prod_{u'' \in L} W(u'') > 1$ [7].

Lemma 3 *With probability $1 - o(1)$ the graph $G(n, d/n)$ satisfies all the conditions (A), (B) and (C).*

The proof of Lemma 3 appears in Section 13.2.

Let $\mathcal{G}_d^A$ denote the family of graphs such that they have no t-core for $t \geq d$ and they satisfy conditions (A), (B) and (C). We show that the chain $\mathcal{M}_c$ has a rapid mixing for any $k \geq \frac{11}{2}d$, as long as the underlying graph $G(n, d/n) \in \mathcal{G}_d^A$. We show rapid mixing by using path coupling in [3]. In particular, Theorem 2 follows as a corollary of Lemma 3 and the following result.

Theorem 6 *Let $\mathcal{M}_c$ be such that the underlying graph $G(n, d/n) \in \mathcal{G}_d^A$ and $k \geq \frac{11}{2}d$. Then, the following is true: Consider X and Y, two states of $\mathcal{M}_c$ such that $H(X, Y) = 1$. Let P_X, P_Y be the transition probabilities of $\mathcal{M}_c$ given that it starts from X, Y, respectively. There is a coupling $\nu_{X,Y}$ of P_X and P_Y such that for (X', Y') distributed as in $\nu_{X,Y}$ it holds that*

$$E_{\nu_{X,Y}}[H(X', Y')] \leq 1 - 0.1225/n, \tag{3}$$

where $E_{\nu_{X,Y}}[\cdot]$ denotes the expectation w.r.t. the distribution $\nu_{X,Y}$.

The proof of Theorem 6 appears in Section 4.

3.2 Hard Core Model - Proof of Theorem 3

As opposed to $\mathcal{M}_c$, the chain $\mathcal{M}_{hc}$ is trivially ergodic. Thus, we proceed by showing that it has rapid mixing. We require from $G(n, d/n)$ to have the following properties: (A) Let $\mathcal{B} = \mathcal{B}(\alpha, \gamma, c)$ such that $\alpha = \epsilon_0/2$, $\gamma = \epsilon_0^2$ and $c = 10$. Each of the blocks in $\mathcal{B}(\alpha, \gamma, c)$ is either a tree or a unicyclic graph. (B) For the values of α, γ, c as in (A) the following holds: For each $B \in \mathcal{B}$ and every u at the outer boundary of B there is no path L from u to a high degree vertex $u' \in B$ such that $\prod_{u'' \in L} W(u'') > 1$.

Corollary 1 *With probability $1 - o(1)$ the graph $G(n, d/n)$ satisfies both the conditions (A) and (B).*

The proof of Corollary 1 is very similar to the proof Lemma 3, so we omit it.

Let $\mathcal{G}_d^B$ be family of graphs that satisfy conditions (A) and (B). For λ as specified in Theorem 3, and underlying graph $G(n, d/n) \in \mathcal{G}_d^B$, we show $\mathcal{M}_{hc}$ has rapid mixing. As in Theorem 2, we use path coupling. In particular, Theorem 3 follows as a corollary of the following theorem and Corollary 1.

Theorem 7 *For $\mathcal{M}_{hc}$ such that $G(n, d/n) \in \mathcal{G}_d^B$ and with fugacity λ as Theorem 3 specifies, the following is true: Consider X and Y, two states of $\mathcal{M}_{hc}$ such that $H(X, Y) = 1$. Let P_X, P_Y be the transition probabilities of $\mathcal{M}_{hc}$ given that it starts from X, Y, respectively. There is a coupling $\nu_{X,Y}$ of P_X and P_Y such that for (X', Y') distributed as in $\nu_{X,Y}$ it holds that*

$$E_{\nu_{X,Y}}[H(X', Y')] \leq 1 - (\epsilon_0 - 20\epsilon_0^2)/n, \tag{4}$$

where $E_{\nu_{X,Y}}[\cdot]$ denotes the expectation w.r.t. the distribution $\nu_{X,Y}$ and $\epsilon_0 = \min\{\epsilon, 1/100\}$.

The proof of Theorem 7 appears in Section 5.

[6] A graph without a t-core can have its vertices ordered $v_1, \ldots v_n$ such that v_i has fewer than t neighbours in $\{v_1, \ldots, v_{i-1}\}$.

[7] For condition (C) there was a discussion in Section 2.

4 Proof of Theorem 6

Consider the chain $\mathcal{M}_c$ with underlying graph $G(n, d/n) \in \mathcal{G}_d^A$ and $k \geq \frac{11}{2}d$. Let X, Y be two states of $\mathcal{M}_c$ at hamming distance 1 from each other, i.e. $H(X, Y) = 1$. Furthermore, assume for the vertex w that we have $X(w) \neq Y(w)$. We are going to show that there is a coupling $\nu_{X,Y}$ of P_X and P_Y that has the property stated in (3). For this, we construct a pair (X', Y') distributed as in $\nu_{X,Y}$ by considering the two copies of $\mathcal{M}_c$ at states X, Y, respectively, and by coupling the next step of these two chains. The coupling is such that we choose the same block to update in both chains.

There are cases where we update the colouring of a single vertex block, i.e. a break-point[8]. There we use the notion of *maximal coupling* transition, as defined in [21]. The update of the colouring of a single vertex u, of degree $(1+\alpha)d$, is made such that the probability of having $X'(u) \neq Y'(u)$ is minimized. This kind of coupling is what we call maximal coupling transition. Let L_X, L_Y be the set of colours not appearing in the neighbourhood of u under the colouring X and Y, respectively. Since we assume $k > (1+\alpha)d$, both sets are not empty. Take two mappings $f_X : [0, 1] \to L_X$ and $f_Y : [0, 1] \to L_Y$ such that

- for each $c \in L_X$, $|f^{-1}(c)| = \frac{1}{|L_X|}$ and similarly for Y,
- $\{x : f_X(x) \neq f_Y(x)\}$ is as small as possible in measure.

Then, take a uniformly random real $\mathcal{U} \in [0, 1]$ and choose colour $f_X(\mathcal{U})$ for $X'(u)$ and $f_Y(\mathcal{U})$ for $Y'(u)$.

Consider now the coupling of the transitions of two chains at state X, Y so as to construction of (X', Y'). Let B be the block whose colouring is chosen to be updated in the coupling. Let $N = |\mathcal{B}|$, i.e. B is chosen with probability $1/N$. We need to consider three cases for the relative position of the block B and the disagreeing vertex w.

`Case 1:` The disagreeing vertex w is internal in a block B', i.e. it is not adjacent to any vertex outside B'. With probability $1/N$ in the coupling we have $B = B'$. Then, we have $H(X', Y') = 0$ as the boundary conditions of B in both chains are identical. Also, with the remaining probability we have $H(X', Y') = 1$. That is

$$E_{\nu_{X,Y}}[H(X', Y')] = 1 - 1/N.$$

`Case 2:` The disagreeing vertex w belongs to a block B' but it is adjacent to some vertices outside B'. Also, assume that B' contains more than one vertices. It is easy to check that $\Delta(w) \leq (1+\alpha)d$ (otherwise condition (C) would be violated). Also, it holds that all the blocks that are adjacent to w are single vertex blocks, i.e. they are break-points.

For this case the following holds: If we don't have $B = B'$ or B adjacent to w, then we have $H(X', Y')$=1. On the other hand, with probability $1/N$ the coupling chooses to update B', i.e. we have $B = B'$. Then, as in Case 1, we have $H(X', Y') = 0$. Also, with probability at most $(1+\alpha)d/N = 1.01d/N$ the block B is a break-point adjacent to w. Then, the probability that the break-point gets different colour assignment in X' and Y' is at most $\frac{1}{k-(1+\alpha)d} = \frac{1}{4.49d}$, as $k \geq \frac{11}{2}d$. That is,

$$E_{\nu_{X,Y}}[H(X', Y')] \leq 1 - 0.75/N.$$

The update of the colouring of u, a neighbour of w, yields to a disagreement with probability at most $\frac{1}{k-(1+\alpha)d}$ for the following reason: We use maximal coupling. Both of the two lists of available colours for u contain at least $k - (1+\alpha)d$ colours and they differ in at most one element. Then it is standard to calculate the upper bound for the probability of creating a disagreement on u.

[8]Break points by definition have small degree, i.e. at most $(1+\alpha)d$.

Case 3: The vertex w is itself a block, i.e. w is a break-point. If we don't have $B = B'$ or B adjacent to w, then we have $H(X', Y')$=1. Otherwise, with probability $1/N$ the coupling chooses to update w, i.e. $B = \{w\}$. Then we have $H(X', Y') = 0$. Also, with probability at most $(1 + \alpha)d/N = 1.01d/N$ the coupling chooses B to be one of the blocks that are adjacent to w. W.l.o.g. assume that B consist of more than one vertex[9]. We need to bound the expected number of disagreements R_B that are generated inside B when its colouring is updated. To this end we use the following proposition.

Proposition 1 *For Case 3, there is a coupling such that* $E[R_B] \leq \frac{0.8688}{d}$.

Using Proposition 1 and the fact that $\alpha = 10^{-2}$ and $N \leq n$, we get that

$$E_{\nu_{X,Y}}[H(X', Y')] \quad \leq \quad 1 - \frac{1}{N} + \frac{(1+\alpha)d}{N} E[R_B] \leq 1 - \frac{0.1225}{n}. \tag{5}$$

The theorem follows.

4.1 Proof of Proposition 1

For the proof of the proposition we use an approach similar to one used in [4] (the proof of Theorem 1 (a)). That is, we use the well-known "disagreement percolation" coupling construction in [2]. We wish to couple $X'(B)$ and $Y'(B)$ as close as possible. However, identical coupling is precluded due to the disagreement at the boundary of B. The disagreement percolation assembles the coupling in a stepwise fashion moving away from w. Disagreements propagates into B along paths from w. A disagreement at vertex $u' \in B$ at distance r from w propagates to a neighbour u at distance $r + 1$ if $X'(u) \neq Y'(u)$. Observing that the distributions of $X'(u), Y'(u)$ are invariant under a Glauber dynamics transition, we may couple $X'(u), Y'(u)$ by using the *maximal coupling*. We get that $Pr[X'(u) \neq Y'(u)] \leq A$, where

$$A = \begin{cases} \frac{2}{k-(1+\alpha)d} & \text{if } \Delta(u) \leq (1+\alpha)d \\ 1 & \text{otherwise.} \end{cases}$$

To see this consider the following:

1. the vertex u has at most 2 neighbours which disagree, since B is either a tree or unicyclic graph,
2. there are at least $k - \min\{k - 1, \Delta(u)\}$ available colours for u.

The disagreement percolation is stochastically dominated by an independent process. This implies that the following is true. Consider a configuration of the vertices in B such that each vertex $u \in B$ is either "disagreeing" or "agreeing". Let the product measure $\mathcal{P}$ which specifies that the vertex u is disagreeing with probability $\rho_u = A$ (where A is as we define above). In this context we call a "path of disagreement" every self avoiding path [10] that has all of its vertices disagreeing.

Let v_0 be the only vertex in B which is adjacent to w. Consider a "agreeing-disagreeing" configuration of the vertices in B acquired according to the measure $\mathcal{P}$. In this configuration, let Z_i be the number of paths of disagreements of length exactly i that start from the vertex v_0. It holds that

$$E[R_B] \leq \sum_{i \geq 0} E_{\mathcal{P}}[Z_i],$$

where $E_{\mathcal{P}}[\cdot]$ denotes the expectation taken w.r.t. the measure $\mathcal{P}$.

Instead of bounding $E_{\mathcal{P}}[Z_i]$ w.r.t B, it is equivalent to study the same quantity on the tree T, the tree of self-avoiding walks defined as follows: T is rooted at vertex v_0, at level i it contains all the vertices in B which are reachable from v_0 through a self-avoiding path inside B of length exactly i. At this point we use the following lemma.

[9] It will be conceivable from the analysis that this is the worst case assumption.

[10] A path L is called *self-avoiding* if there are no two $u_j, u_{j'} \in L$ such that $u_j = u_{j'}$.

Lemma 4 *Consider a tree $H = (V, E)$ containing vertices of different degrees. Let l_i denote the set of vertices at the level i in H. For a vertex $v \in H$ let L_v denote the simple path connecting it to the root r. For L_v we define the following weight:*

$$\mathcal{C}_{p,s}(L_v) = \prod_{u \in L_v} \left(\mathbb{I}_{\{\Delta(u) \le s\}} \cdot p + \mathbb{I}_{\{\Delta(u) > s\}} \right),$$

where $p \in [0,1]$ and $s > 0$ is an integer. Assume that for any vertex v such that $\Delta(v) > s$ the following condition holds:

$$\prod_{u \in L_v} \left(\frac{\mathbb{I}_{\{\Delta(u) \le s\}}}{(1+\gamma)} + d^{10} \cdot \Delta(u) \cdot \mathbb{I}_{\{\Delta(u) > s\}} \right) \le (1+\gamma), \tag{6}$$

for some $\gamma > 0$. Assume that d, s are sufficiently large, also $d, s \gg (1+\gamma)$ while both $(s \cdot p), (d \cdot p) \in [\frac{1}{100(1+\gamma)}, \frac{1}{1+\gamma}]$. Then (6) implies that

$$\sum_{v \in l_i} \mathcal{C}_{p,s}(L_v) \le p \cdot (1-\theta)^i \qquad \forall i \ge 0, \tag{7}$$

for any $\theta \le \min\{1 - ps(1+\gamma), 1 - (ps)^{9/10}\}$.

The proof of Lemma 4 appears in Section 4.2.

At this point, observe the following: Let L_v be a path in T from the root r to some vertex v. The probability that L_v is a path of disagreement is upper bounded by the quantity $\mathcal{C}_{p,s}(L_v)$ where $p = \frac{2}{\frac{11}{2}d - (1+\alpha)d} = \frac{1}{2.245d}$, $s = (1+\alpha)d = 1.01d$. Thus, it holds that

$$E_{\mathcal{P}}[Z_i] \le \sum_{v \in l_i} \mathcal{C}_{p,s}(L_v),$$

where l_i is the set of vertices at level i in T. Observe, also, that the quantities p, d and s we consider here satisfy the restrictions imposed in Lemma 4 about their (relative) size. Finally, each L_v in T satisfies the condition in (6) as the disagreeing vertex w is a break-point and the root of T is adjacent to w. Thus, Lemma 4 implies that $E_{\mathcal{P}}[Z_i] \le p(1-\theta)^i$. So as to bound $E_{\mathcal{P}}[Z_i]$ we take the maximum possible value for θ, w.r.t. d, s, p. It is direct that $\theta \le 0.5127$. Thus, it holds that

$$E_{\mathcal{P}}[Z_i] \le \frac{0.44543}{d} (0.4873)^i.$$

From the above we get that $E[R_B] \le \frac{0.44543}{d} \cdot \frac{1}{0.5127} = \frac{0.8688}{d}$, as promised.

4.2 Proof of Lemma 4

Proof of Lemma 4: Assume that we are given γ, s, p. When there is no danger of confusion, we abbreviate $\mathcal{C}_{p,s}$ to $\mathcal{C}$. Also let $w_1, \ldots, w_{\Delta(r)}$ denote the children of r in H and let H_j be the subtree of H rooted at the child w_j.

From the condition in (6) it is direct to see that the root r is of degree at most s. Then, it is trivial to verify that (7) holds, i.e. (7) is true for $i = 0$. Consider now that $i > 0$. Since $\Delta(r) \le s$, it holds that

$$\sum_{v \in l_i} \mathcal{C}^H(L_v) = p \cdot \sum_{j=1}^{\Delta(r)} \left(\sum_{v \in l_{i-1}} \mathcal{C}^{H_j}(L_v) \right), \tag{8}$$

where the superscripts H, H_j over $\mathcal{C}$ denote the tree w.r.t. which we consider the paths and the weights $\mathcal{C}(L_v)$. Since $\Delta(r) \leq s$, a sufficient condition for (7) to hold is the following one: For each H_j it should hold that

$$\sum_{v \in l_{i-1}} \mathcal{C}^{H_j}(L_v) \leq p(1-\theta)^i \times \frac{1}{ps}, \tag{9}$$

i.e. if the bound in (9) holds, then we can plug it into (8) and get (7).

Observe that in the previous step we used the fact that $\Delta(r) \leq s$ to "unfold" the condition in (7). We can repeating the same step on the subtrees of H_j. This gives rise to a weighting over each path L_v from r to $v \in l_i$. We denote this weight as $R(L_v)$. It is not hard to see that for L_v we get that

$$\mathcal{R}(L_v) = p(1-\theta)^i \times \left(\frac{1}{ps}\right)^{i-q} \times \prod_{u \in M \setminus \{v\}} \frac{1}{\Delta(u)} \times \left(\mathbb{I}_{\{\Delta(v) \leq s\}} \frac{1}{p} + \mathbb{I}_{\{\Delta(v) > s\}} \right), \tag{10}$$

where the set M contains all the vertices in L_v of degree larger than s while $|M \setminus \{v\}| = q$. Furthermore, it is direct to see that (7) holds as long as for every path L_v between r and $v \in l_i$ we have that

$$\mathcal{R}(L_v) \geq 1. \tag{11}$$

The lemma will follow by showing that (11) holds. We consider three cases, depending on the degree of v, the last vertex in L_v and the cardinality of the set M: (A) $\Delta(v) \leq s$, (B) $\Delta(v) > s$ and $|M \setminus \{v\}| = 0$, (C) $\Delta(v) > s$ and $|M \setminus \{v\}| \geq 1$. We show the following results.

Claim 1 *For the case (A), the condition in (11) is true as long as* $\theta \leq 1 - ps(1+\gamma)$.

Claim 2 *For the case (B), the condition in (11) is true as long as* $\theta \leq 1 - (ps)^{9/10}$.

Claim 3 *For the case (C), the condition in (11) is true as long as* $\theta \leq 1 - ps(1+\gamma)$.

Clearly, condition in (11) holds for $\theta \leq \min\{1 - ps(1+\gamma), 1 - (ps)^{9/10}\}$. The lemma follows. $\diamond$

Proof of Claim 1: Consider case (A). For each path L_v, the condition (6) implies that

$$\frac{1}{(1+\gamma)^{i-q}} \prod_{u \in M} d^{10} \cdot \Delta(u) \leq (1+\gamma).$$

If $q \geq 1$, we have that

$$\begin{aligned} \mathcal{R}(L_v) &= p(1-\theta)^i \frac{1}{(ps)^{i-q}} \frac{1}{p} \prod_{u \in M \setminus \{v\}} (\Delta(u))^{-1} \geq (1-\theta)^i \left(\frac{1}{ps(1+\gamma)}\right)^{i-q} \frac{d^{10q}}{1+\gamma} \\ &\geq \left(\frac{1-\theta}{ps(1+\gamma)}\right)^i \frac{\left[d^{10} ps(1+\gamma)\right]^q}{1+\gamma} \geq \left(\frac{1-\theta}{ps(1+\gamma)}\right)^i, \end{aligned}$$

where the last inequality follows from the fact for large d we have that $\frac{\left[d^{10}ps(1+\gamma)\right]^q}{1+\gamma} \gg 1$.

If $q = 0$, then it is direct to see that $\mathcal{R}(L_v) = \left(\frac{1-\theta}{ps}\right)^i$. In any case, the claim follows. $\diamond$

Proof of Claim 2: Given the assumptions of the case (B), the condition in (6) implies that $\frac{1}{(1+\gamma)^i}d^{10}\cdot\Delta(v)\leq 1+\gamma$. In turn, this implies that $i\geq\frac{\ln[d^{10}\Delta(v)]}{\ln(1+\gamma)}-1$. Then, we have that

$$\mathcal{R}(L_v) \;=\; p(1-\theta)^i\frac{1}{(ps)^i}\geq\frac{p}{(ps)^{i/10}}\cdot\left(\frac{1-\theta}{(ps)^{9/10}}\right)^i.$$

The claim will follow by showing that $(ps)^{i/10}<p$. Observe that we have that $ps<1$ and $i\geq\frac{\ln[d^{10}\Delta(v)]}{\ln(1+\gamma)}-1$. That is,

$$\begin{aligned}(ps)^{i/10} &\leq (ps)^{\frac{1}{10}\frac{\ln(d^{10}\Delta(v))}{\ln(1+\gamma)}-\frac{1}{10}}\leq(d^{10}\Delta(v))^{-\frac{1}{10}\frac{\ln(ps)}{\ln\left(\frac{1}{1+\gamma}\right)}}\cdot(ps)^{-1/10}\\ &\leq (d^{10}\Delta(v))^{-\frac{1}{10}}(ps)^{-1/10} \qquad\qquad [\text{since } ps\leq 1/(1+\gamma)]\\ &\leq \frac{1}{d(\Delta(v))^{1/20}}\left(\frac{(ps)^{-1}}{\Delta(v)^{1/2}}\right)^{1/10}\leq p.\end{aligned}$$

The last inequality follows by taking large d,s ($\Delta(v)>s$). The claim follows. $\diamond$

Proof of Claim 3: For the case (C), the condition in (6) implies that

$$\frac{1}{(1+\gamma)^{i-q}}\prod_{u\in M\backslash\{v\}}d^{10}\cdot\Delta(u)\leq 1+\gamma.$$

where $q\geq 1$, i.e. $M\backslash\{v\}\neq\emptyset$. Due to our assumptions about d, p, s and γ, we made in the statement of Lemma 4 it holds that $\frac{p}{1+\gamma}[d^{10}ps(1+\gamma)]^q\gg 1$, for large d. Thus, we have that

$$\begin{aligned}\mathcal{R}(L_v) &= p(1-\theta)^i\frac{1}{(ps)^{i-q}}\prod_{u\in M\backslash\{v\}}(\Delta(u))^{-1}\geq\frac{p}{1+\gamma}(1-\theta)^i\left(\frac{1}{ps(1+\gamma)}\right)^{i-q}d^{10q}\\ &\geq \frac{p}{1+\gamma}(1-\theta)^i\left(\frac{1}{ps(1+\gamma)}\right)^{i}[d^{10}ps(1+\gamma)]^q\\ &\geq \left(\frac{1-\theta}{ps(1+\gamma)}\right)^i. \qquad\qquad \left[\text{since } \frac{p}{1+\gamma}(d^{10}ps(1+\gamma))^q\gg 1\right]\end{aligned}$$

The claim follows. $\diamond$

5 Proof of Theorem 7

Consider the chain $\mathcal{M}_{hc}$ with underlying graph $G(n,d/n)\in\mathcal{G}_d^B$ and λ as in the statement of Theorem 3. Let X,Y be two states of $\mathcal{M}_{hc}$ at hamming distance 1 from each other, i.e. $H(X,Y)=1$. Furthermore, assume for the vertex w that we have $X(w)\neq Y(w)$. We are going to show that there is a coupling $\nu_{X,Y}$ of P_X and P_Y that has the property stated in (4). For this, we construct a pair (X',Y') distributed as in $\nu_{X,Y}$ by considering the two copies of $\mathcal{M}_c$ at states X,Y, respectively, and by coupling the next step of these two chains. The coupling is such that we choose the same block to update in both chains.

Let B be the block whose configuration is chosen to be updated in the coupling. Let $N=|\mathcal{B}|$, i.e. B is chosen with probability $1/N$. We need to consider three cases for the relative position of the block B and the disagreeing vertex w. The cases we consider are exactly the same as those we considered for the colouring model in the proof of Theorem 6.

`Case 1:` The disagreeing vertex w is an internal vertex of a block B', i.e. it is not adjacent to any vertex outside B'. It is direct to see with probability $1/N$ in the coupling $B = B'$. Then, we have $H(X', Y') = 0$ as the boundary conditions of B in both chains are identical. Also, with the remaining probability we have $H(X', Y') = 1$. That is

$$E_{\nu_{X,Y}}[H(X', Y')] = 1 - 1/N.$$

`Case 2:` The disagreeing vertex w belongs to a block B' but it is adjacent to some vertices outside B'. Assume, also, that B' contains more than one vertices. It is easy to check that $\Delta(w) \leq (1 + \alpha)d$ (otherwise condition (B) would be violated). Also, it holds that all the blocks that are adjacent to w are single vertex blocks, i.e. they are break-points.

For this case the following holds: If we don't have $B = B'$ or B adjacent to w, then we have $H(X', Y')$=1. Otherwise, with probability $1/N$ the coupling chooses to update B', i.e. we have $B = B'$. Then, as in Case 1, we have $H(X', Y') = 0$. Also, with probability at most $(1 + \alpha)d/N = (1 + \frac{\epsilon_0}{2})d/N$ the block B is a break-point adjacent to w. Then, the probability that the break-point becomes disagreeing is at most $\frac{\lambda}{1+\lambda} \leq \frac{1-\epsilon_0}{2d} + \frac{1}{2d^2}$. Then it is direct to get that

$$E_{\nu_{X,Y}}[H(X', Y')] \leq 1 - 1/(2N).$$

The update of the spin of u, a neighbour of w, yields to a disagreement with probability at most $\frac{\lambda}{1+\lambda}$ for the following reason: W.l.o.g. assume that $X(w)$ is *occupied*, i.e. w belongs to the independent set and $Y(w)$ is *unoccupied*. Clearly $X(u)$ cannot be occupied. The only way we can have disagreement is when all the neighbours of u, apart from w, in both configurations are unoccupied. Then, $Y(u)$ becomes occupied (disagreeing) with probability $\frac{\lambda}{1+\lambda}$.

`Case 3:` The vertex w is itself a block, i.e. w is a break-point. If we don't have $B = B'$ or B adjacent to w, then we have $H(X', Y')$=1. Otherwise, with probability $1/N$ the coupling chooses to update w, i.e. $B = \{w\}$. Then we have $H(X', Y') = 0$. Also, with probability at most $(1 + \alpha)d/N$ the coupling chooses B to be one of the blocks that are adjacent to w. W.l.o.g. assume that all these blocks consist of more than one vertex[11]. We need to bound the expected number of disagreements R_B that are generated inside B from the update. To this end we use the following result.

Proposition 2 *For Case 3, there is a coupling such that* $E[R_B] \leq \left(1 - \frac{3}{2}\epsilon_0 + 20\epsilon_0^2\right) \frac{1}{d} + \frac{10}{d^2}$.

Using Proposition 2 and the fact that $N \leq n$ we get that

$$E_{\nu_{X,Y}}[H(X', Y')] \leq 1 - \frac{1}{N} + \frac{(1 + \alpha)d}{N} E[R_B] \leq 1 - \frac{\epsilon_0 - 20\epsilon_0^2}{n}.$$

The theorem follows.

5.1 Proof of Proposition 2

The proof of this proposition is very similar to the proof of Proposition 1 for the colouring model. That is, we use the "disagreement percolation" coupling construction from citeDisPerc. We consider configurations such that each vertex $u \in B$ is either "disagreeing" or "agreeing". We let the product measure $\mathcal{P}$ specify that the vertex u is disagreeing with some appropriate probability ρ_u.

We consider a "agreeing-disagreeing" configuration of the vertices in B acquired according to the measure $\mathcal{P}$. In this configuration, let Z_i be the number of path of disagreements of length exactly i that

[11] It will be conceivable from the analysis that this is the worst case assumption.

start from the vertex v_0. The vertex $v_0 \in B$ is the only vertex which is adjacent to the disagreeing vertex w. It holds that

$$E[R_B] \leq \sum_{i \geq 0} E_{\mathcal{P}}[Z_i],$$

where $E_{\mathcal{P}}[\cdot]$ denotes the expectation taken w.r.t. the measure $\mathcal{P}$.

Compared to the case of colourings, the only difference we have here is the value of the probability ρ_v in the disagreement percolation. Observe that for every $u \in B$, the distributions of $X'(u), Y'(u)$ are invariant under a Glauber dynamics transition. We can couple $X'(u), Y'(u)$ by using the *maximal coupling*. Working as in the case of a single vertex update case of Theorem 7, we get that $\rho_u = \frac{\lambda}{1+\lambda}$.

Having specified the measure $\mathcal{P}$ we work as follows: Instead of bounding $E_{\mathcal{P}}[Z_i]$ w.r.t B, it is equivalent to study the same quantity on the tree T, the tree of self-avoiding walks we defined in the proof of Proposition 1. At this point we use the following lemma.

Lemma 5 *Consider a tree $H = (V, E)$ containing vertices of different degrees. Let l_i denote the set of vertices at the level i in H. For a vertex $v \in H$ let L_v denote the simple path connecting it to the root r. For L_v we define the following weight:*

$$\mathcal{C}_{p,s}(L_v) = p^{|L_v|},$$

where $p \in [0, 1]$. Given $d, s > 0$, for any vertex v such that $\Delta(v) > s$ the following condition holds:

$$\prod_{u \in L_v} \left(\frac{\mathbb{I}_{\{\Delta(u) \leq s\}}}{(1+\gamma)} + d^{10} \cdot \Delta(u) \cdot \mathbb{I}_{\{\Delta(u) > s\}} \right) \leq (1+\gamma), \tag{12}$$

for some $\gamma > 0$. Assume that d, s are sufficiently large and $d, s \gg (1+\gamma)$, while both $(s \cdot p)$, $(d \cdot p) \in [\frac{1}{100} \frac{1}{1+\gamma}, \frac{1}{1+\gamma}]$. Then (12) implies that

$$\sum_{v \in l_i} \mathcal{C}_{p,s}(L_v) \leq p \cdot (1-\theta)^i \qquad \forall i \geq 0, \tag{13}$$

for any $\theta \leq 1 - ps(1+\gamma)$.

The proof of Lemma 5 appears in Section 5.2.

Let $\lambda_0 = \frac{1-\epsilon_0}{2d}$ for fixed $\epsilon_0 = \min\{\epsilon, 1/100\}$, and let d be sufficiently large. Let L_v be a path in T from the root r to some vertex $v \in l_i$, where l_i the vertices at level i in T. The probability that L_v is a path of disagreement is upper bounded by the quantity $\mathcal{C}_{p,s}(L_v)$ where $p = \frac{\lambda_0}{1+\lambda_0}$, $s = (1+\epsilon_0/2)d$ and let $\gamma = \epsilon_0^2$. Thus, it holds that

$$E_{\mathcal{P}}[Z_i] \leq \sum_{v \in l_i} \mathcal{C}_{p,s}(L_v).$$

Observe that the quantities p, d and s we consider here satisfy the restrictions imposed in Lemma 5 about their (relative) size. Finally, L_v satisfies the condition in (12) as the disagreeing vertex w is a break-point and the root of T is adjacent to w. Thus, Lemma 5 implies that $E_{\mathcal{P}}[Z_i] \leq p(1-\theta)^i$. So as to bound $E_{\mathcal{P}}[Z_i]$ we take the maximum possible value for θ, w.r.t. d, s, p. According to Lemma 5 it should hold that $\theta \leq 1 - ps(1+\gamma) = \frac{1+\epsilon_0/2-6\epsilon_0^2}{2}$. We get that

$$E[R_B] = \frac{p}{\theta} \leq \left(1 - \frac{3}{2}\epsilon_0 + 20\epsilon_0^2\right) \frac{1}{d} + \frac{10}{d^2}.$$

The proposition follows.

5.2 Proof of Lemma 5

Proof of Lemma 5: When there is no danger of confusion, we abbreviate $\mathcal{C}_{p,s}$ to $\mathcal{C}$. Also let $w_1, \ldots, w_{\Delta(r)}$ denote the children of r in H and let H_j be the subtree of H rooted at the child w_j.

From the condition in (12) it is direct to see that the root r is of degree at most s. Then, it is trivial to verify that (13) holds, i.e. (13) is true for $i = 0$. Consider now that $i > 0$. Since $\Delta(r) \leq s$, it holds that

$$\sum_{v \in l_i} \mathcal{C}^H(L_v) = p \cdot \sum_{j=1}^{\Delta(r)} \left(\sum_{v \in l_{i-1}} \mathcal{C}^{H_j}(L_v) \right), \tag{14}$$

where the superscripts H, H_j over $\mathcal{C}$ denote the tree w.r.t. which we consider the paths and the weights $\mathcal{C}(L_v)$. Since $\Delta(r) \leq s$, a sufficient condition for (13) to hold is the following one: For each H_j it should hold that

$$\sum_{v \in l_{i-1}} \mathcal{C}^{H_j}(L_v) \leq p(1-\theta)^i \times \frac{1}{ps}, \tag{15}$$

i.e. if the bound in (15) holds, then we can plug it into (14) and get (13).

Observe that in the previous step we used the fact that $\Delta(r) \leq s$ to "unfold" the condition in (13). Repeating the same step on the subtrees of H_j gives rise to a weighting over each path L_v from r to $v \in l_i$. Let us denote this weight as $R(L_v)$. It is not hard to see that for L_v we get that

$$\mathcal{R}(L_v) = p(1-\theta)^i \times \left(\frac{1}{ps}\right)^{i-q} \times \prod_{u \in M \backslash \{v\}} \frac{1}{\Delta(u)} \times \frac{1}{p}, \tag{16}$$

where M contains all the vertices in L_v of degree larger than s while $|M \backslash \{v\}| = q$. Furthermore, it is direct to see that (13) holds as long as for every path L_v between r and $v \in l_i$ we have that

$$\mathcal{R}(L_v) \geq 1. \tag{17}$$

It suffice to show that (17) holds. Observe that, for each path L_v, the condition (12) implies that

$$\frac{1}{(1+\gamma)^{i-q}} \prod_{u \in M} d^{10} \cdot \Delta(u) \leq (1+\gamma).$$

If $q \geq 1$, we have that

$$\begin{aligned}
\mathcal{R}(L_v) &= p(1-\theta)^i \frac{1}{(ps)^{i-q}} \frac{1}{p} \prod_{u \in M \backslash \{v\}} (\Delta(u))^{-1} \geq (1-\theta)^i \left(\frac{1}{ps(1+\gamma)}\right)^{i-q} \frac{d^{10q}}{1+\gamma} \\
&\geq \left(\frac{1-\theta}{ps(1+\gamma)}\right)^i \frac{\left[d^{10} ps(1+\gamma)\right]^q}{1+\gamma} \geq \left(\frac{1-\theta}{ps(1+\gamma)}\right)^i,
\end{aligned}$$

where the last inequality follows from the fact that $\frac{\left[d^{10} ps(1+\gamma)\right]^q}{1+\gamma} \gg 1$ for large d.

If $q = 0$, then it is direct to see that $\mathcal{R}(L_v) = \left(\frac{1-\theta}{ps}\right)^i$. Clearly, in any case it should hold that $\theta \leq 1 - ps(1+\gamma)$. The lemma follows. $\diamond$

6 Tail Bounds

6.1 A Tail bound for weighting schema 1

Consider the setting where we assign the vertices of $G(n, d/n)$ weights as in Section 1.3, in (2). That is, consider a graph $G_{n,d/n}$ and fixed numbers $\alpha, \gamma, c > 0$. Consider a path $L = v_1, \ldots, v_{|L|}$ in $G(n, d/n)$. For a vertex $v_i \in L$ we let the quantity $W(v_i)$ be such that

$$W(v_i) = \begin{cases} (1+\gamma)^{-1} & \text{if } \Delta(v_i) \leq (1+\alpha)d \\ d^c \cdot \Delta(v_i) & \text{otherwise}, \end{cases}$$

where $\Delta(v_i)$ is the degree of v_i. Also, let

$$C(L) = \prod_{v \in L} W(v).$$

A very useful result for the analysis that follows is the following theorem.

Theorem 8 *Let P be a path in $G(n, d/n)$ with a number of vertices $|P| \leq 100 \ln n$. For fixed $\alpha \in (0, 3/2)$, $\gamma > 0, c > 0$ and any $\delta > 0$ it holds that*

$$Pr[C(P) \geq \delta] \leq \exp\left[-d^{4/5}\left(|P| + \ln \delta\right)\right].$$

The proof of Theorem 8 appears in Section 10.

6.2 A tail bound for weighting schema 2

We consider the weighting schema we are going to use in Section 8. That is, consider a path L in $G(n, d/n)$ and for each $v_i \in L$ we let N_i denote the set of the vertices outside L which are adjacent to v_i, i.e. $|N_i| = \Delta(v_i) - 2$. For every $w \in N_i$ let $E_{out}(w)$ denote the influence [12] on vertex w, only from paths of length at most $\ln n/d^{2/5}$ that do not use vertices in L. For every v_i we let

$$Q(v_i) = \max_{w \in N_i} \{E_{out}(w)\}.$$

Now, we associate each vertex in L with the following quantity

$$U(v_i) = \begin{cases} \frac{\max\{1, Q(v_i)\}}{1+\gamma} & \text{if } \Delta(v_i) \leq (1+\alpha)d \\ \max\{1, Q(v_i)\} \cdot d^c \cdot \Delta(v_i) & \text{otherwise.} \end{cases}$$

Also, we let

$$AC(L) = \prod_{v_i \in L} U(u_i).$$

Also, we remind the reader that a path L is called *elementary* if there is no cycle of length less than $10 \ln n/d^{2/5}$ that contains two distinct vertices of the path.

Theorem 9 *Let P be an* elementary *path in $G(n, d/n)$ with a number of vertices $|P| \leq \frac{\ln n}{\ln^2 d}$. For fixed $\alpha \in (0, 3/2)$, $\gamma > 0, c > 0$ and any $\delta > 0$ it holds that*

$$Pr[AC(P) \geq \delta] \leq 2\exp\left[-d^{7/10}\left(|P| + \ln \delta\right)\right].$$

The proof of Theorem 9 appears in Section 12.

[12]For a definition of the concept influence, see in Section 2, Definition 1

7 Proof of Theorem 5

For proving Theorem 5, (i.e. in Proposition 3) we need to use the tail bounds from Section 6.

Let $L = v_1, \ldots, v_{|L|}$ be an *elementary* path in $G(n, d/n)$ of length $T = \frac{\ln n}{\ln^5 d}$. Also, let ϱ_L be the probability that L does not have a break point. It holds that

$$\begin{aligned} Pr[\mathbb{U} \neq \emptyset] &\leq E[|\mathbb{U}|] \leq \binom{n}{T+1} \cdot \left(\frac{d}{n}\right)^T \cdot \varrho_L && \text{[by linearity of expectation]} \\ &\leq \left(\frac{ne}{T+1}\right)^{T+1} \cdot \left(\frac{d}{n}\right)^T \cdot \varrho_L && \text{[as } \tbinom{n}{i} \leq (ne/i)^i\text{]} \\ &\leq nd^T \cdot \varrho_L \leq n^{1.1} \cdot \varrho_L. && \text{[as } T = \tfrac{\ln n}{\ln^5 d}\text{]} \end{aligned} \tag{18}$$

Thus, it suffices to bound appropriately ϱ_L. Consider the weighting schema in in Section 1.3, in (2).

Definition 3 *A vertex $v_i \in L$ is called* left-break *or* right-break *for L if it has the corresponding property below:*

right-break: *There is no path $L' \in \mathcal{P}(v)$ such that $\prod_{v_i \in L'} W(v_i) > 1$ and $L' \cap L$ contains v_j for $j \leq i$,*

left-break: *There is no path $L' \in \mathcal{P}(v)$ such that $\prod_{v_i \in L'} W(v_i) > 1$ and $L' \cap L$ contains v_j for $j \geq i$.*

(For the definition of the set $\mathcal{P}(u)$ see in Definition1.)

Remark. Clearly, $v_i \in L$ is break point *if and only if* it is both a left-break and a right-break for L.

Proposition 3 *Consider the elementary path L of length of $\frac{\ln n}{\ln^5 d}$ in $G(n, d/n)$. Let Y_l and Y_r be the number of left breaks and the number of right breaks for L, respectively. Let the event*

$$\mathcal{S} = \text{"either } Y_l \text{ or } Y_r \text{ is smaller than } 0.9|L|\text{"}.$$

For α, γ, c as in the statement of Theorem 5 and sufficiently large d, it holds that

$$Pr[\mathcal{S}] \leq 8n^{-\frac{1}{2}\frac{\gamma}{1+\gamma}\ln d}.$$

The proof of Proposition 3 appears in Section 8.

Given that $Y_l, Y_r \geq 0.9|L|$, it is easy to see that the number of vertices in L which are at the same time left-break and right-break is at least $0.8|L|$. Thus, we get that

$$\varrho_L \leq Pr[\mathcal{S}] \leq 8n^{-\frac{1}{2}\frac{\gamma}{1+\gamma}\ln d}.$$

The theorem follows by plugging the above inequality into (18).

8 Proof of Proposition 3

It is not hard to see that the random variables Y_l and Y_r are symmetric, i.e. identically distributed. We will focus only Y_r and we will show that:

$$Pr[Y_r < 0.9|L|] \leq 4n^{-\frac{1}{2}\frac{\gamma}{1+\gamma}\ln d}. \tag{19}$$

Then the proposition will follow by a simple union bound, i.e.

$$Pr[\mathcal{S}] \leq Pr[Y_r < 0.9|L|] + Pr[Y_l < 0.9|L|] \leq 2Pr[Y_r < 0.9|L|],$$

where the last inequality follows from symmetry.

For each $v_i \in L$ let N_i denote the set of the vertices outside L which are adjacent to v_i, i.e. $|N_i| = \Delta(v_i) - 2$. For every $w \in N_i$ let $E_{out}(w)$ denote the influence on vertex w, only from paths of length at most $\ln n/d^{2/5}$ that do not use vertices in L. For every v_i let

$$Q(v_i) = \max_{w \in N_i} \{E_{out}(w)\}.$$

Now, we associate each vertex in L with the following quantity

$$U(v_i) = \begin{cases} \frac{\max\{1, Q(v_i)\}}{1+\gamma} & \text{if } \Delta(v_i) \le (1+\alpha)d \\ \max\{1, Q(v_i)\} \cdot d^c \cdot \Delta(v_i) & \text{otherwise,} \end{cases}$$

where α, γ, c are defined in the statement of the proposition.

Remark. The reader should observe that due to the assumption that L is an elementary path the weight of each vertex in L is independent of the weight of the other vertices in L.

It is straightforward that any vertex v_i such that $U(v_i) > 1$ cannot be a right break for L. Let

$$H = \{v_j \in L | U(v_j) > 1\}. \tag{20}$$

For each $v_j \in H$ let $\mathcal{R}_j = v_j, v_{j+1}, \ldots, v_s$ be the maximal subpath of L such that for any $j' \in [j, s]$ it holds $\prod_{r=j}^{j'} U(v_r) > 1$.

Lemma 6 *A vertex $v_{j'} \notin H$ is a right break if there is no $v_j \in H$ such that $j \le j'$ and $v_{j'} \in \mathcal{R}_j$.*

The proof of Lemma 6 appears in Section 13.4. From Lemma 6 we get the following: Letting $\mathcal{R} = \cup_j \mathcal{R}_j$, it holds that $Y_r = |L| - |\mathcal{R}|$. The proposition will follow by deriving an appropriate tail bound for $|\mathcal{R}|$. That is, it suffice to show that

$$Pr\left[|\mathcal{R}| \ge 0.1|L|\right] \le 2n^{-\frac{1}{2}\frac{\gamma}{1+\gamma}\ln d}. \tag{21}$$

To this end, we work as follows: Let $\mathbb{A}$ be the event that "the maximum degree in $G(n, d/n)$ is at most $\ln^2 n$". Let, also, $\mathbb{F}$ be the event "H has cardinality less than $\frac{\ln d}{d^{7/10}} \ln n$". For any $t > 0$, we have that

$$\begin{aligned} Pr[|\mathcal{R}| \ge |L|/10] &\le Pr[\mathbb{A}^c] + Pr[\mathbb{F}^c] + Pr[|\mathcal{R}| \ge |L|/10 | \mathbb{A}, \mathbb{F}] \\ &\le Pr[\mathbb{A}^c] + Pr[\mathbb{F}^c] + Pr[e^{t|\mathcal{R}|} \ge e^{t|L|/10} | \mathbb{A}, \mathbb{F}] \\ &\le Pr[\mathbb{A}^c] + Pr[\mathbb{F}^c] + \frac{E\left[e^{t|\mathcal{R}|} | \mathbb{A}, \mathbb{F}\right]}{\exp(t|L|/10)}, \end{aligned} \tag{22}$$

where in the last inequality we used Markov's inequality. We need to bound $Pr[\mathbb{A}^c]$, $Pr[\mathbb{F}^c]$ and $E\left[e^{t|\mathcal{R}|} | \mathbb{A}, \mathbb{F}\right]$. As far as $Pr[\mathbb{F}^c]$ is regarded, we use the following lemma.

Lemma 7 *For α, γ, c as in the statement of Proposition 3, it holds that*

$$Pr\left[|H| \ge \frac{\ln d}{d^{7/10}} \ln n\right] \le n^{-\frac{1}{2}\frac{\gamma}{1+\gamma}\ln d}.$$

The proof of Lemma 7 appears in Section 8.1. As far as $Pr[\mathbb{A}^c]$ is regarded we get that

$$Pr[\mathbb{A}^c] \le n Pr[\Delta(v) > \ln^2 n] \le \exp\left(-\ln^2 n\right). \tag{23}$$

In the first inequality, above, we use a union bound (over all vertices), while the second inequality follows from Chernoff's bound. As far as $E\left[e^{t|\mathcal{R}|} | \mathbb{A}, \mathbb{F}\right]$ is regarded, we use the following proposition.

Proposition 4 *For the path L and α, γ, c as in the statement of Proposition 3 the following is true: For any $0 \leq t \leq d^{3/5}$ it holds that*

$$E\left[e^{t|\mathcal{R}|}|\mathbb{A}, \mathbb{F}\right] \leq \exp\left(2d^{3/10} \ln d \cdot \ln n\right).$$

The proof of Proposition 4 appears in Section 9. Plugging into (22) the bounds from Proposition 4, Lemma 7, (23) and setting $t = d^{3/5}$ we get (21), as promised. The proposition follows.

8.1 Proof of Lemma 7

First we are going to compute $Pr[v_i \in H] = Pr[U(v_i) > 1]$. Proposition 5 (in Section 10) implies that $E\left[U^t(v_i)|\mathbb{A}\right] \leq 2\exp\left(-\frac{\gamma}{1+\gamma}t\right)$, for any $1 \leq t < d^{7/10}$. Then, for $t = d^{7/10}$ we have that

$$\begin{aligned} Pr[U(v_i) > 1|\mathbb{A}] &= Pr[U^t(v_i) > 1|\mathbb{A}] \leq E\left[U^t(v_i)|\mathbb{A}\right] \\ &\leq 2\exp\left(-\frac{\gamma}{1+\gamma}d^{7/10}\right). \end{aligned}$$

Let $N = \frac{\ln d}{d^{7/10}} \ln n$. If $|H| \geq N$, then there should be a subset B of vertices in L such that $|B| = N$ and $B \subseteq H$. It holds that

$$\begin{aligned} Pr[|H| > N] &\leq \binom{\frac{\ln n}{\ln^5 d}}{N} \exp\left(-\frac{\gamma}{1+\gamma}d^{7/10}\right)^N \\ &\leq n^{\frac{\ln 2}{\ln^5 d}} \exp\left(-\frac{\gamma}{1+\gamma}d^{7/10}N\right) & \left[\text{as } \binom{n_0}{N} \leq 2^{n_0}\right] \\ &\leq n^{\frac{\ln 2}{\ln^5 d}} \cdot n^{-\frac{\gamma}{1+\gamma}\ln d} \leq n^{-\frac{1}{2}\frac{\gamma}{1+\gamma}\ln d}, \end{aligned}$$

where the last inequality holds for large d. The lemma follows.

9 Proof of Proposition 4

Proposition 4 follows as a corollary from the following two lemmas.

Lemma 8 *For a path L and α, γ, c and t as in the statement of Proposition 4 the following is true:*

$$E\left[e^{t|\mathcal{R}|}|\mathbb{A}, \mathbb{F}\right] \leq \left(E\left[e^{t|\mathcal{R}_1|}|\mathbb{A}, \mathbb{F}, \mathbb{B}\right]\right)^{\frac{\ln d}{d^{7/10}}\ln n},$$

where the event $\mathbb{B}$ = "$v_1 \in H$" and the set H is defined in (20).

The proof of Lemma 8 appears in Section 9.1. Also, we have the following result.

Lemma 9 *For a path L and α, γ, c, t as in the statement of Proposition 4 the following is true:*

$$E\left[e^{t|\mathcal{R}_1|}|\mathbb{A}, \mathbb{F}, \mathbb{B}\right] \leq e^{2d}.$$

The proof of Lemma 9 appears in Section 9.2.

9.1 Proof of Lemma 8

We denote the expectation operator $E[\cdot|\mathbb{A},\mathbb{F}]$ as $E_{\mathbb{A},\mathbb{F}}$ as well as for $Pr[\cdot|\mathbb{A},\mathbb{F}]$ we use $P_{\mathbb{A},\mathbb{F}}[\cdot]$.

Assume that we have $j < j'$ such that $\mathcal{R}_j \cap \mathcal{R}_{j'} \neq \emptyset$. Also, let $\mathcal{R}_{j'} = w_{j'}, \ldots, w_s$. Then, it holds that $\mathcal{R}_j = w_j, \ldots, w_s$ as $\mathcal{R}_j, \mathcal{R}_{j'}$ should have the same rightmost vertex (this follows by the definition of these sets). Let the set of positive integers

$$J = \left\{ j \in \mathbb{N} | v_j \in H, \text{ and } \nexists j' < j : \mathcal{R}_j \subset \mathcal{R}_{j'} \right\}.$$

It holds that $|\mathcal{R}| = \sum_{j \in J} |\mathcal{R}_j|$. Let j_1 denote the minimum element in J, j_2 the second smallest and so on. Clearly it should hold that $j_1, j_2, j_3 \leq |L|$. In what follows, by conditioning that $j_2 = |L| + 1$ we imply that the set J has at most one element, i.e. j_1.

$$\begin{aligned}
E_{\mathbb{A},\mathbb{F}}[\exp(t|\mathcal{R}|)] &= E_{\mathbb{A},\mathbb{F}}\left[\prod_{j \in J} \exp(t|\mathcal{R}_j|)\right] \\
&= \sum_{s=1}^{|L|+1} E_{\mathbb{A},\mathbb{F}}\left[\prod_{j \in J} \exp(t|\mathcal{R}_j|) \,|j_2 = s\right] P_{\mathbb{A},\mathbb{F}}[j_2 = s] \\
&= \sum_{s=1}^{|L|+1} E_{\mathbb{A},\mathbb{F}}\left[\prod_{j \in J \setminus \{j_1\}} \exp(t|\mathcal{R}_j|) \,|j_2 = s\right] E[\exp(t\mathcal{R}_{j_1})|j_2 = s] P_{\mathbb{A},\mathbb{F}}[j_2 = s] \\
&\leq E_{\mathbb{A},\mathbb{F}}[\exp(t|\mathcal{R}_{j_1}|)] \cdot \max_s \left\{ E_{\mathbb{A},\mathbb{F}}\left[\prod_{j \in J \setminus \{j_1\}} \exp(t|\mathcal{R}_j|) \,|j_2 = s\right]\right\}.
\end{aligned}$$

Observe that the above factorization holds since the underlying path L we consider is elementary. That is there are not two different vertices $v_i, v_j \in L$ whose weights are dependent with each other. Assume that the above expectation is maximized for $s = s_0$. It holds that

$$E_{\mathbb{A},\mathbb{F}}[\exp(t|\mathcal{R}|)] \leq E_{\mathbb{A},\mathbb{F}}[\exp(t|\mathcal{R}_{j_1}|)] \cdot E_{\mathbb{A},\mathbb{F}}\left[\prod_{j \in J \setminus \{j_1\}} \exp(t|\mathcal{R}_j|) \,|A_1\right],$$

where $A_1 =$ "$j_2 = s_0$". Working in the same manner we get that

$$E_{\mathbb{A},\mathbb{F}}[\exp(t|\mathcal{R}|)] \quad \leq \quad \prod_{s=0}^{T} E_{\mathbb{A},\mathbb{F}}[\exp(t|\mathcal{R}_{j_s}|) \mid \cap_{q=1}^{s} A_q],$$

where $T = \frac{\ln d}{d^{7/10}} \ln n - 1$, since we have assumed that $|H| \leq \frac{\ln d}{d^{7/10}} \ln n$. Also, the events A_q are defined in the same manner as the event A_1.

Consider now the event A_q. Observe that $\mathcal{R}_j$s are correlated with each other in the sense that $|\mathcal{R}| \leq |L|$. Revealing some of them affects *only* the starting point of the rest, i.e. the events $\cap_{q=1}^{s} A_q$ specify the leftmost position that $\mathcal{R}_{j_{s+1}}$ can start. This is due to the fact that since the underlying path is elementary and the weights $U(v_i)$s of the vertices in the path are independent. The further $\mathcal{R}_{j_{s+1}}$ starts the less available vertices it has to extent. E.g. it may be the case that $v_{j_{s+1}}$ is the last vertex of the path and there are no extra vertices to extend $\mathcal{R}_{j_{s+1}}$. Considering all the above, it is direct that for any $s \in J$ the following holds:

$$E_{\mathbb{A},\mathbb{F}}[\exp(t\mathcal{R}_{j_s})| \cap_{q=1}^{s} A_q] \leq E_{\mathbb{A},\mathbb{F}}[\exp(t|\mathcal{R}_1|)|\mathbb{B}],$$

where the event $\mathbb{B} =$"$v_1 \in H$" and the set H is defined in (20). The lemma follows.

9.2 Proof of Lemma 9.

Let $Pr[\cdot|\mathbb{A},\mathbb{F},\mathbb{B}]$ be denoted as $P_{\mathbb{A},\mathbb{F},\mathbb{B}}[\cdot]$. It holds that

$$\begin{aligned} E\left[e^{t|\mathcal{R}_1|}|\mathbb{A},\mathbb{F},\mathbb{B}\right] &= \sum_{r=1}^{|L|} e^{tr}\cdot P_{\mathbb{A},\mathbb{F},\mathbb{B}}[|\mathcal{R}_1|=r] \leq \sum_{r=1}^{|L|} e^{tr}\cdot P_{\mathbb{A},\mathbb{F},\mathbb{B}}[|\mathcal{R}_1|\geq r] \\ &\leq \sum_{r=1}^{|L|} e^{tr}\cdot P_{\mathbb{A},\mathbb{F},\mathbb{B}}[AC(L_{1,r})>1], \end{aligned} \tag{24}$$

where $L_{1,r}$ denotes the subpath of L which starts from vertex v_1 and ends at the vertex v_r. It holds that

$$\begin{aligned} Pr[AC(L_{1,r})>1|\mathbb{A},\mathbb{F},\mathbb{B}] &= \frac{Pr[AC(L_{1,r})>1,\mathbb{B}|\mathbb{A},\mathbb{F}]}{Pr[\mathbb{B}|\mathbb{A},\mathbb{F}]} \\ &\leq \frac{Pr[AC(L_{1,r})>1|\mathbb{A},\mathbb{F}]}{Pr[\mathbb{B}|\mathbb{A},\mathbb{F}]}. \end{aligned} \tag{25}$$

For the nominator we use the tail bound we obtained in Theorem 9 (in Section 6.2). Then, we get that

$$Pr[AC(L_{1,r})>1|\mathbb{A},\mathbb{F}] \leq 4\exp\left(-d^{7/10}r\right). \tag{26}$$

Also, it holds that

$$Pr[\mathbb{B}|\mathbb{A},\mathbb{F}] \geq Pr[\Delta(v_1)\geq(1+\alpha)d] \geq Pr[\Delta(v_1)=(1+\alpha)d] \geq \exp(-2d). \tag{27}$$

The last inequality follows from Stirling's approximation and the fact that $\alpha\in(0,3/2)$. Plugging (26) and (27) into (25) we get that

$$Pr[AC(L_{1,r})>1|\mathbb{A},\mathbb{F},\mathbb{B}] \leq \exp\left[-d^{7/10}r+2d\right].$$

In turn, plugging the above into (24) we have that

$$\begin{aligned} E\left[e^{t|\mathcal{R}_1|}|\mathbb{A},\mathbb{F},\mathbb{B}\right] &\leq 4e^{2d}\sum_{r=1}^{|L|}\exp\left[(t-d^{7/10})r\right] \leq 4e^{2d}\sum_{r=1}^{|L|}\exp\left[-0.8d^{7/10}r\right] && [\text{as } 0\leq t\leq d^{3/5}] \\ &\leq 4e^{2d}\frac{\exp\left[-0.8d^{7/10}\right]}{1-\exp\left[-0.8d^{7/10}\right]} \leq e^{2d}. && \left[\text{as } e^{0.8d^{0.7}}/(1-e^{0.8d^{0.7}})\leq 1/4\right] \end{aligned}$$

The lemma follows.

10 Proof of Theorem 8

Let the $\mathbb{A}$ be the event that "the maximum degree in $G(n,d/n)$ is at most $\ln^2 n$". It holds that

$$Pr[C(P)\geq\delta] \leq Pr[\mathbb{A}^c] + Pr[C(P)\geq\delta|\mathbb{A}]. \tag{28}$$

From the union bound we get that

$$Pr[\mathbb{A}^c] \leq nPr\left[\text{degree}(v)>\ln^2 n\right] \leq n\exp\left(-1.2\ln^2 n\right) \leq \exp\left(-\ln^2 n\right), \tag{29}$$

where in the second inequality we use Chernoff's bound. Consider, now, the quantity

$$K(P)=\ln C(P)=\sum_{u\in L}\ln W(u)=\sum_{u\in L}R(u),$$

where $R(u) = \ln W(u)$. The proposition will follow by bounding appropriately the probability $Pr[K(L) > \ln \delta]$. For a real $t > 0$ it holds that

$$\begin{aligned} Pr[C(P) \geq \delta|\mathbb{A}] &= Pr[K(P) \geq \ln \delta|\mathbb{A}] = Pr[\exp(tK(L)) \geq \delta^t|\mathbb{A}] \\ &\leq E\left[\exp(tK(P))|\mathbb{A}\right]/\delta^t, \end{aligned} \tag{30}$$

where the last inequality follows from Markov's inequality. For the calculation of $E\left[\exp(tK(P))|\mathbb{A}\right]$, we use the following lemma.

Lemma 10 *Consider the path $P = v_1, \ldots, v_{|P|}$, where $|P| \leq 100 \ln n$. It holds that*

$$E[\exp(tK(P))|\mathbb{A}] \leq \prod_{i=1}^{|P|} E[\exp(tR(v_i))|\mathbb{A}, B_i],$$

where the event $B_i =$"the number of edges between the set $V_i = \{v_1, \ldots, v_{i-1}\}$ and the set $V \backslash V_i$ is equal to $(i-1)(\ln^2 n - 2)$ and the maximum degree in V_i is at most $\ln^2 n$".

The proof of Lemma 10 appears in Section 10.1.

Proposition 5 *For fixed $C > 0$, an integer t such that $0 < t < C \cdot d^{4/5}$ and any $v_i \in P$ the following is true: For α, γ, c as in the statement of Theorem 8 it holds that*

$$E\left[\exp(t \cdot R(v_i))|\mathbb{A}, B_i\right] \leq \exp\left(-\frac{\gamma}{1+\gamma}t\right) + 3d^{t(1+c)} \exp\left(-\frac{\alpha^2}{5}d\right),$$

where the event B_i is defined in the statement of Lemma 10.

The proof of Proposition 5 appears in Section 11. Using Lemma 10 and Proposition 5 we get that

$$\begin{aligned} E[\exp(tK(P))|\mathbb{A}] &\leq \left[\exp\left(-\frac{\gamma}{1+\gamma}t\right) + 3d^{t(1+c)} \exp\left(-\frac{\alpha^2}{5}d\right)\right]^{|P|} \\ &\leq \exp\left(-\frac{\gamma}{1+\gamma}t|P|\right)\left[1 + 3d^{t(1+c)} \exp\left(-\frac{\alpha^2}{5}d + \frac{\gamma}{1+\gamma}t\right)\right]^{|P|} \\ &\leq \exp\left(-\frac{\gamma}{1+\gamma}t|P| + 3d^{t(1+c)} \exp\left(-\frac{\alpha^2}{5}d + \frac{\gamma}{1+\gamma}t\right)|P|\right), \end{aligned}$$

where in the last inequality we used the fact that $1 + x \leq e^x$. Taking $t = 5\frac{1+\gamma}{\gamma}d^{4/5}$ and sufficiently large d we get that

$$E[\exp(tK(P))|\mathbb{A}] \leq \exp\left(-d^{4/5}|P|\right).$$

The proposition follows by plugging the above inequality into (30).

10.1 Proof of Lemma 10

Proof of Lemma 10: It is elementary to verify that

$$E\left[\exp(tK(P))|\mathbb{A}\right] = \sum_{j=1}^{\ln^2 -1} E\left[\exp(tK(P))|\mathbb{A}, \Delta(v_1) = j\right] Pr[\Delta(v_1) = j|\mathbb{A}]. \tag{31}$$

Also, it is direct that

$$
\begin{aligned}
E\left[\exp\left(tK(P)\right)|\mathbb{A},\Delta_{v_1}=j\right] &= E\left[\exp\left(tR(v_1)\right)|\mathbb{A},\Delta_{v_1}=j\right]\cdot E\left[\prod_{s=2}^{|P|}\exp\left(tR(v_s)\right)|\mathbb{A},\Delta_{v_1}=j\right] \\
&\leq E\left[\exp\left(tR(v_1)\right)|\mathbb{A},\Delta_{v_1}=j\right]E\left[\prod_{s=2}^{|P|}\exp\left(tR(v_s)\right)|\mathbb{A},\Delta_{v_1}=\ln^2 n\right].
\end{aligned}
$$

The above inequality holds since $R(v_s)$s are increasing with $\Delta(v_1)$. Plugging the above inequality into (31) we get that

$$
E\left[\exp\left(tK(P)\right)|\mathbb{A}\right]\leq E\left[\exp\left(tR(v_1)\right)|\mathbb{A}\right]\cdot E\left[\prod_{s=2}^{|P|}\exp\left(tR(v_s)\right)|\mathbb{A},\Delta(v_1)=\ln^2 n\right].
$$

The lemma follows by working in the same manner of the remaining vertices in the path. $\diamond$

11 Proof of Proposition 5

It is direct that $\exp\left(t\cdot R(v_i)\right)=W^t(v_i)$, i.e. $E[\exp(tR(v_i))|\mathbb{A},B_i]=E[W^t(v_i)|\mathbb{A},B_i]$.

We distinguish the edges which are incident to v_i into three categories: The first one consists of the edges between v_i and the vertices in the set $\{v_1,\ldots,v_{i-2}\}$. Let $\Delta^{int}(v_i)$ be the number of these edges. The second category consists of the edges between v_i and the vertices not in $\{v_1,\ldots,v_{i+1}\}$. Let $\Delta^{ext}(v_i)$ be the number of these edges. The third is the edges that v_i used to connect to v_{i-1} and v_{i+1}. Of course if $i=|P|$ then there is no $v_{|P|+1}$. I.e. it holds that

$$
\Delta(v_i)\leq\Delta^{int}(v_i)+\Delta^{ext}(v_i)+2.
$$

Conditional on B_i, the marginal of $\Delta^{int}(v_i)$ follows the binomial distribution with parameters $(i-2)(\ln^2 n-2)$, $1/(n-i+1)$. Also, the marginal of $\Delta^{ext}(v_i)$ follows the binomial distribution with parameters $n-(i+1)$, d/n. The joint distribution of $\Delta^{int}(v_i),\Delta^{ext}(v_i)$ is such that the degree of v_i is at most equal to the maximum degree in $G(n,d/n)$. It holds that

$$
\begin{aligned}
E[W^t(v_i)|\mathbb{A},B_i] &= \sum_{j=0}^{i\ln^2 n}E[W^t(v_i)|\mathbb{A},B_i,\Delta^{int}(v_i)=j]Pr[\Delta^{int}(v_i)=j|\mathbb{A},B_i] \\
&\leq E[W^t(v_i)|\mathbb{A},B_i,\Delta^{int}_{v_i}=0]+\sum_{j=1}^{i\ln^2 n}E[W^t(v_i)|\mathbb{A},B_i,\Delta^{int}_{v_i}=j]Pr[\Delta^{int}_{v_i}=j|\mathbb{A},B_i],
\end{aligned}
$$

where we used the fact that $Pr[\Delta^{int}_{v_i}=0|\mathbb{A},B_i]\leq 1$. Also, it holds that

$$
Pr[\Delta^{int}_{v_i}=j|\mathbb{A},B_i]\leq\frac{Pr[\Delta^{int}_{v_i}=j|B_i]}{Pr[\mathbb{A}|B_i]}\leq\frac{Pr[\Delta^{int}_{v_i}=1|B_i]}{Pr[\mathbb{A}|B_i]}. \tag{32}
$$

The last inequality following from the following observation. Given the event B_i, $\Delta^{int}_{v_i}$ is binomially distributed. The function $f(j)=Pr[\Delta^{int}_{v_i}=j|B_i]$ is decreasing for any $j>E[\Delta^{int}_{v_i}|B_i]$. As $E[\Delta^{int}_{v_i}|B_i]\leq\ln^3 n/n$, $f(j)$ is decreasing for every $j\geq 1$. That is for any $j\geq 1$ it holds that $Pr[\Delta^{int}_{v_i}=j|B_i]\leq Pr[\Delta^{int}_{v_i}=1|B_i]$. Also, it is trivial to show that $Pr[\mathbb{A}|B_i]\geq 1/2$.

Using this fact, the above inequality can be written as follows:

$$\begin{aligned}
E[W^t(v_i)|\mathbb{A},B_i] &\le E[W^t(v_i)|\mathbb{A},B_i,\Delta^{int}_{v_i}=0] + 2Pr[\Delta^{int}_{v_i}=1|B_i]\cdot\sum_{j=1}^{i\ln^2 n} E[W^t(v_i)|\mathbb{A},B_i,\Delta^{int}_{v_i}=j] \\
&\le E[W^t(v_i)|\mathbb{A},B_i,\Delta^{int}_{v_i}=0] + 200\frac{\ln^3 n}{n}\sum_{j=1}^{i\ln^2 n} E[W^t(v_i)|\mathbb{A},B_i,\Delta^{int}_{v_i}=j], \qquad (33)
\end{aligned}$$

as $i \le 100\ln n$. The proposition will follow by bounding appropriately $E[W^t(v_i)|\mathbb{A},B_i,\Delta^{int}_{v_i}=0]$ and $E[W^t(v_i))|\mathbb{A},B_i,\Delta^{int}_{v_i}=j]$. For this, we provide the following two lemmas.

Lemma 11 *For α,γ,c,t as in the statement of Proposition 5, it holds that*

$$E[W^t(v_i)|\mathbb{A},B_i,\Delta^{int}_{v_i}=0] \le \exp\left(-\frac{\gamma}{1+\gamma}t\right) + 2d^{t(c+1)}\exp\left(-\frac{\alpha^2}{5}d\right).$$

The proof of Lemma 11 appears in Section 11.1. As far as $E[W^t(v_i)|\mathbb{A},B_i,\Delta^{int}_{v_i}=j]$ is regarded we use the following lemma.

Lemma 12 *For integer j such that $1 \le j \le \ln^4 n - 2$ and α,γ,c,t as in the statement of Proposition 5, it holds that*

$$E[W^t(v_i)|\mathbb{A},B_i,\Delta^{int}_{v_i}=j] \le 5(2d^c)^t(\ln n)^{4t}.$$

The proof of Lemma 12 appears in Section 11.2. Plugging into (33) the bounds from Lemmas 11, 12 we get that

$$\begin{aligned}
E[W^t(v_i)|\mathbb{A},B_i] &\le \exp\left(-\frac{\gamma}{1+\gamma}t\right) + 2d^{t(t+c)}\exp\left(-\frac{\alpha^2}{4}d\right) + 1000(2d)^t\frac{(\ln n)^{4t+6}}{n} \\
&\le \exp\left(-\frac{\gamma}{1+\gamma}t\right) + 3d^{t(c+1)}\exp\left(-\frac{\alpha^2}{4}d\right). \qquad [\text{as } t \le Cd^{4/5}]
\end{aligned}$$

The proposition follows.

11.1 Proof of Lemma 11

Proof of Lemma 11: So as to prove the lemma we need the following claim.

Claim 4 *For sufficiently large d, integers q,t such that $|q|,|t| \le Cd^{4/5}$, for any constant $C>0$, and $0<\alpha<3/2$ the following is true:*

$$\sum_{i=(1+\alpha)d+q}^{n}(i+2)^t\binom{n}{i}\left(\frac{d}{n}\right)^i\left(1-\frac{d}{n}\right)^{n-i} \le d^t\cdot\exp\left(-\frac{\alpha^2}{5}d\right).$$

Letting $S_A = E[W^t(v_i)|\mathbb{A},B_i,\Delta^{int}_{v_i}=0]$ we have the following:

$$\begin{aligned}
S_A &= (1+\gamma)^{-t}Pr[\Delta^{ext}(v_i) \le (1+\alpha)d-2|\mathbb{A},B_i] + d^{ct}\sum_{i=(1+\alpha)d-2}^{n}(i+2)^t Pr[\Delta^{ext}(v_i)=i|\mathbb{A},B_i] \\
&\le (1+\gamma)^{-t} + \frac{d^{ct}}{Pr[\mathbb{A}]}\sum_{i=(1+\alpha)d-2}^{n}(i+2)^t\binom{n}{i}\left(\frac{d}{n}\right)^i\left(1-\frac{d}{n}\right)^{n-i} \quad [\text{as } Pr[\Delta^{ext}_{v_i} \le (1+\alpha)d-2|\mathbb{A},B_i] \le 1] \\
&\le (1+\gamma)^{-t} + \frac{d^{t(c+1)}}{Pr[\mathbb{A}]}\exp\left(-\frac{\alpha^2}{4}d\right) \quad [\text{from Claim 4}] \\
&\le (1+\gamma)^{-t} + 2d^{t(c+1)}\exp\left(-\frac{\alpha^2}{4}d\right). \quad [\text{as } Pr[\mathbb{A}] \ge 1/2]
\end{aligned}$$

The lemma follows by noting that $(1+\gamma)^{-t} \leq \exp\left(-\frac{\gamma}{1+\gamma}t\right)$. $\diamond$

Proof of Claim 4: Let $S = \sum_{i=(1+\alpha)d+q}^{n}(i+2)^t\binom{n}{i}\left(\frac{d}{n}\right)^i\left(1-\frac{d}{n}\right)^{n-i}$. Using the fact that $\binom{n}{i} = \frac{n}{i}\binom{n-1}{i-1}$, for $1 \leq i \leq n$, we get that

$$\begin{aligned}
S &= d\sum_{i=(1+\alpha)d+q}^{n}\left(1+\frac{2}{i}\right)^t i^{t-1}\binom{n-1}{i-1}\left(\frac{d}{n}\right)^{i-1}\left(1-\frac{d}{n}\right)^{n-1-(i-1)} \\
&\leq d\sum_{j=(1+\alpha)d+q-1}^{n-1}\left(1+\frac{2}{j+1}\right)^t (j+1)^{t-1}\binom{n-1}{j}\left(\frac{d}{n}\right)^{j}\left(1-\frac{d}{n}\right)^{n-1-j}. \qquad \text{[we set } j=i-1\text{]}
\end{aligned}$$

It is direct to see that repeating the exactly the same calculation l times, where $l \leq t$, and for sufficiently large d we get that

$$S \leq d^l\left(\prod_{s=0}^{l-1}\left(1+\frac{s}{d}\right)\right)\cdot\sum_{j=(1+\alpha)d+q-l}^{n-l}\left(1+\frac{2}{j+l}\right)^t (j+l)^{t-l}\binom{n-l}{j}\left(\frac{d}{n}\right)^{j}\left(1-\frac{d}{n}\right)^{n-l-j}.$$

Also, for $l=t$ we see that the above inequality is equivalent to the following

$$S \leq d^t\cdot e^2\cdot Pr[\mathcal{B}(n-t,d/n) > (1+\alpha)d+q-t]\cdot\prod_{s=0}^{t-1}\left(1+\frac{s}{d}\right), \tag{34}$$

where the last term expresses the probability that a random variable distributed as in binomial distribution with parameters $n-t$ and d/n is at least $(1+\alpha)d+q-t$. Using the fact that $|q|,|t| < Cd^{4/5}$ and standard large deviation results about the binomial distribution, i.e. Corollary 2.3 from [14] we get that

$$Pr[\mathcal{B}(n-t,d/n) \geq (1+\alpha)d+q-t] \leq \exp\left(-\alpha^2 d/4\right). \tag{35}$$

Also, we have that $\prod_{s=0}^{t-1}\left(1+\frac{s}{d}\right) \leq \exp\left(\sum_{s=0}^{t-1}s/d\right) \leq \exp\left(t(t-1)/(2d)\right)$. Taking large d and using the fact that $t \leq Cd^{4/5}$ we get that $\exp\left(t(t-1)/(2d)\right) \leq \exp\left(C^2 d^{3/5}/2\right)$. Plugging this result and (35) into (34) we get that $S \leq d^t\exp\left(-\frac{\alpha^2}{5}d\right)$, as promised. $\diamond$

11.2 Proof of Lemma 12

Proof of Lemma 12: For $S_{B,j} = E\left[\exp\left(t\cdot R(v_i)\right)|\mathbb{A}, B_i, \Delta^{int}(v_i)=j\right]$, where $1 \leq j \leq \ln^4 n - 2$, we work as follows:

$$\begin{aligned}
S_{B,j} &\leq (1+\gamma)^{-1} + \frac{d^{ct}}{Pr[\mathbb{A}]}\sum_{i=\max\{(1+\alpha)d-j-2,0\}}^{\ln^2 n}(i+j+2)^t\binom{n}{i}\left(\frac{d}{n}\right)^i\left(1-\frac{d}{n}\right)^{n-i} \\
&\leq (1+\gamma)^{-1} + \frac{d^{ct}}{Pr[\mathbb{A}]}\sum_{i=0}^{\ln^2 n}(i+\ln^4 n)^t\binom{n}{i}\left(\frac{d}{n}\right)^i\left(1-\frac{d}{n}\right)^{n-i} \qquad [\text{as } 1 \leq j \leq \ln^4 n - 2] \\
&\leq (1+\gamma)^{-1} + 2d^{ct}(\ln n)^{4t}\sum_{i=0}^{\ln^2 n}(1+i/\ln^4 n)^t\binom{n}{i}\left(\frac{d}{n}\right)^i\left(1-\frac{d}{n}\right)^{n-i} \\
&\leq (1+\gamma)^{-1} + 2d^{ct}(\ln n)^{4t}\sum_{i=0}^{\ln^2 n}2^t\binom{n}{i}\left(\frac{d}{n}\right)^i\left(1-\frac{d}{n}\right)^{n-i} \qquad [\text{as } 1+i/\ln^4 n \leq 2 \text{ for } i \leq \ln^4 n] \\
&\leq (1+\gamma)^{-1} + 2(2d^c)^t(\ln n)^{4t}\cdot Pr[\mathcal{B}(n,d/n) \leq \ln^2 n] \\
&\leq 5(2d^c)^t(\ln n)^{4t}.
\end{aligned}$$

The lemma follows. $\diamond$

12 Poof of Theorem 9

Let $\mathbb{A}$ be the event that "the maximum degree in $G(n,d/n)$ is at most $\ln^2 n$". We have that

$$Pr[AC(P)\geq \delta] \leq Pr[\mathbb{A}^c] + Pr[AC(P)\geq \delta|\mathbb{A}], \tag{36}$$

from the law of total probability. Also, from the union bound we get that

$$Pr[\mathbb{A}^c] \leq nPr\left[\text{degree}(v) > \ln^2 n\right] \leq \exp\left(-\ln^2 n\right). \tag{37}$$

where the last inequality follows from Chernoff's bound. Consider, now, the quantity

$$AK(P) = \ln AC(P) = \sum_{v_i\in P} \ln U(v_i) = \sum_{v_i\in P} F(v_i),$$

where $F(v_i) = \ln U(v_i)$. The proposition will follow by bounding appropriately the probability $Pr[AK(L) > \ln\delta]$. For a real $t>0$ it holds that

$$\begin{aligned} Pr[AC(P)\geq \delta|\mathbb{A}] &= Pr[AK(P)\geq \ln\delta|\mathbb{A}] = Pr[\exp\left(t\cdot AK(L)\right)\geq \delta^t|\mathbb{A}] \\ &\leq E\left[\exp\left(t\cdot AK(P)\right)|\mathbb{A}\right]/\delta^t, \end{aligned} \tag{38}$$

where the last inequality follows from Markov's inequality. Due to our assumption that P is elementary the random variables $F(v_1),\ldots,F(v_{|P|})$ are independent with each other. Thus, we have that

$$E[\exp(t\cdot AK(P))|\mathbb{A}] = \prod_{i=1}^{|P|} E[\exp(t\cdot F(v_i))|\mathbb{A}]. \tag{39}$$

Proposition 6 *Assume that α,γ,c are as in the statement of Theorem 9. For a fixed $C>0$ and integer t such that $0<t<Cd^{7/10}$ and any $v_i\in P$ the following it true:*

$$E\left[\exp\left(t\cdot F(v_i)\right)|\mathbb{A}\right] \leq \exp\left(-\frac{\gamma}{1+\gamma}t\right)\left(1+2\exp\left(-\frac{d^{4/5}}{2}\right)\right).$$

The proof of Proposition 6 appears in Section 12.1. Using (39) and Proposition 6 we get that

$$E[\exp(t\cdot AK(P))|\mathbb{A}] \leq \exp\left(-\frac{\gamma}{1+\gamma}t|P| + 2\exp\left(-\frac{d^{4/5}}{2}\right)|P|\right).$$

For $t = 5\frac{1+\gamma}{\gamma}d^{7/10}$ and sufficiently large d we get that

$$E[\exp(t\cdot AK(P))|\mathbb{A}] \leq \exp\left(-d^{7/10}|P|\right). \tag{40}$$

The proposition follows by combining (40), (38) (37) and (36).

12.1 Proof of Proposition 6

It is direct that $\exp(t\cdot F(v_i)) = U^t(v_i)$, i.e. $E[\exp(tF(v_i))|\mathbb{A}] = E[U^t(v_i)|\mathbb{A}]$.

For a vertex $v_i \in P$, let $\delta(v_i)$ denote the number of edges between v_i and the vertices outside the path P. Observe that the $\delta(v_i)$ is stochastically dominated by the binomial distribution with parameters, n, d/n, e.g. $\mathcal{B}(n, d/n)$. By definition it holds that

$$\begin{aligned} E[U^t(v_i)|\mathbb{A}] &\leq \sum_{j=0}^{\ln^2 n-2} E[U^t(v_i)|\mathbb{A}, \delta(v_i)=j]Pr[\delta(v_i)=j|\mathbb{A}] \\ &\leq \sum_{j=0}^{\ln^2 n-2} E[W^t(v_i)|\mathbb{A}, \delta(v_i)=j]\cdot E[Q^t(v_i)|\mathbb{A}, \delta(v_i)=j]\cdot Pr[\delta(v_i)=j|\mathbb{A}], \quad (41) \end{aligned}$$

where $W(v_i)$ is the quantity defined in (2). Observe that the random variables $W(v_i)$ and $Q(v_i)$ are independent with each other given $\delta(v_i)$. For this reason it holds that $E[U^t(v_i)|\Delta(v_i)] = E[W^t(v_i)|\Delta(v_i)]\cdot E[F^t(v_i)|\Delta(v_i)]$. We are going to bound these two expectations by using the following lemmas.

Lemma 13 *For t, α, γ, c as in the statement of Proposition 6, it holds that*

$$E[Q^t(v_i)|\mathbb{A}, \delta(v_i)] < 1 + 2d\cdot\delta(v_i)\exp\left(-d^{4/5}\right). \quad (42)$$

The proof of Lemma 13 appears in Section 12.2. Plugging (42) into (41), we get that

$$E[U^t(v_i)|\mathbb{A}] \leq E[W^t(v_i)|\mathbb{A}] + 2d\exp\left(-d^{4/5}\right)\sum_{j=0}^{\ln^2 n} jE[W^t(v_i)|\mathbb{A}, \delta(v_i)=j]Pr[\delta(v_i)=j|\mathbb{A}]. \quad (43)$$

Lemma 14 *For t, α, γ, c as in the statement of Proposition 6, it holds that*

$$\sum_{j=0}^{\ln^2 n} jE[W^t(v_i)|\mathbb{A}, \delta(v_i)=j]Pr[\delta(v_i)=j|\mathbb{A}] \leq 2d(1+\gamma)^{-t} + 2d^{t(c+1)+1}\exp\left(-\alpha^2 d/5\right).$$

The proof of Lemma 14 appears in Section 12.3.

Lemma 15 *For t, α, γ, c as in the statement of Proposition 6, it holds that*

$$E[W^t(v_i)|\mathbb{A}] \leq \exp\left(-\frac{\gamma}{1+\gamma}t\right) + 2d^{t(c+1)}\exp\left(-\frac{\alpha^2}{5}d\right).$$

We do not provide any proof for Lemma 15 as it is the same as the proof of Lemma 11 (which appears in Section 11.1). Plugging into (43) the bounds from Lemmas 14, 15 we get that

$$\begin{aligned} E[U^t(v_i)|\mathbb{A}] &\leq \left[\exp\left(-\frac{\gamma}{1+\gamma}t\right) + 2d^{t(c+1)}\exp\left(-\frac{\alpha^2}{5}d\right)\right]\left(1+\exp\left(-\frac{d^{4/5}}{2}\right)\right) \\ &\leq \exp\left(-\frac{\gamma}{1+\gamma}t\right)\left(1+2\exp\left(-\frac{d^{4/5}}{2}\right)\right). \qquad \text{[as } t\leq d^{7/10}\text{]} \end{aligned}$$

The proposition follows.

12.2 Proof of Lemma 13

Proof of Lemma 13: For $\delta(v_i) = 0$ it is direct that $Q(v_i) = 1$. Thus, the inequality (42) holds for $\delta(v_i) = 0$. In what follows we consider that $\delta(v_i) > 0$.

Let $w_1, \ldots, w_{\delta(v_i)}$ be the neighbours of v_i outside the path P. Also, for w_i let S_i be the set of paths of length $\ln n/d^{4/5}$ that start from w_i but do not use the vertices in P. For $1 \leq j \leq \delta(v_i)$, let $\mathcal{E}_{j,x}$ be the event that there is a path $L \in S_j$ such that $C(L) \geq x$. Observe that the events $\mathcal{E}_{j,x}$ are identically distributed as j varies. Using Theorem 8 we get the following:

Claim 5 *For $1 \leq q \leq \delta(v_i)$ and any real $x \geq 1$, it holds that*

$$Pr[\mathcal{E}_{q,x}] \leq 2d \cdot \exp\left(-d^{4/5}(\ln x + 1)\right).$$

The proof of Claim 5 follows after this proof. We have that

$$\begin{aligned} E[Q^t(v_i)|\delta(v_i)] &= Pr[Q(v_i) < 1|\delta(v_i)] + E[Q^t(v_i)|\delta(v_i), Q(v_i) \geq 1] \cdot Pr[Q(v_i) \geq 1|\delta(v_i)] \\ &\leq 1 + E[Q^t(v_i)|\delta(v_i), Q(v_i) \geq 1] \cdot Pr[Q(v_i) \geq 1|\delta(v_i)], \end{aligned} \tag{44}$$

as $Pr[Q(v_i) < 1|\delta(v_i)] \leq 1$. Also, it holds that

$$\begin{aligned} \mathcal{S}_A &= E[Q^t(v_i)|\delta(v_i), Q(v_i) \geq 1] \cdot Pr[Q(v_i) \geq 1|\delta(v_i)] \\ &\leq Pr[Q(v_i) \geq 1|\delta(v_i)] \cdot \int_1^{\exp[\ln^2 n]} x^t Pr[Q(v_i) \geq x|\delta(v_i), Q(v_i) \geq 1] dx \\ &\leq \int_1^{\exp[\ln^2 n]} x^t Pr[Q(v_i) \geq x|\delta(v_i)] dx. \end{aligned}$$

The bound $\exp(\ln^2 n)$ follows from simple calculations which suggest that, given that the maximum degree is $\ln^2 n$, for any path $L \in S_j$ it holds that $C(L) \leq \exp\left(\ln^2 n\right)$. Note that

$$\begin{aligned} Pr[Q(v_i) \geq x|\Delta(v_i)] &= Pr[\cup_{i=1}^{\delta(v_i)} \mathcal{E}_{i,x}|\Delta(v_i)] \\ &\leq \Delta(v_i) Pr[\mathcal{E}_{1,x}] && \text{[from the union bound]} \\ &\leq 2d\Delta(v_i) \exp\left(-d^{4/5}(1 + \ln x)\right). && \text{[from Claim 5]} \end{aligned}$$

Thus, we get that

$$\begin{aligned} \mathcal{S}_A &\leq 2d\Delta(v_i) \exp(-d^{4/5}) \int_1^{\exp[\ln^2 n]} x^t \exp\left(-d^{4/5} \ln x\right) dx \\ &\leq 2d\Delta(v_i) \exp(-d^{4/5}) \int_1^{\exp[\ln^2 n]} \exp\left(-\frac{1}{2} d^{4/5} \ln x\right) dx && \text{[as } 0 \leq t \leq Cd^{7/10}\text{]} \\ &\leq 2d\Delta(v_i) \exp(-d^{4/5}). \end{aligned}$$

where in the final derivation we used the fact that $\int_1^{e^{\ln^2 n}} x^{-d^{4/5}/2} \leq 1$. The lemma follows by plugging the above bound for S_A into (44). ◇

Proof of Claim 5: As in the proof of Lemma 13 we use the following terminology: Let $w_1, \ldots, w_{\delta(v_i)}$ be the neighbours of v_i outside the path P. Also, for w_i let S_i be the set of paths that start from w_i but do not use the vertices in P and they are of length at most $\ln n/d^{4/5}$.

For the event $\mathcal{E}_{1,x}$ to occur, there should be a path $L \in S_i$ such that $C(L) \geq x$. Let $S_i^l \subseteq S_i$ denote the paths in S_i of l vertices. It holds that $1 \leq l \leq \ln n/d^{4/5}$. For every $L \in S_i$ let I_L be an indicator variable such that $I_L = 1$ if $C(L) \geq x$, otherwise $I_L = 0$. Applying Markov's inequality we get that

$$Pr[\mathcal{E}_{1,x}] = Pr\left[\sum_{L \in S_i} I_L > 0\right] \leq E\left[\sum_{L \in S_i} I_L > 0\right]. \tag{45}$$

From Theorem 8 we get the following: For any $L \in S_i^j$ we have that

$$Pr[I_L] \leq \exp\left[-d^{4/5}(j + \ln x)\right]. \tag{46}$$

Thus, by the linearity of expectation, it holds that

$$\begin{aligned}
E\left[\sum_{L \in S_i} I_L > 0\right] &= \sum_{j=1}^{\ln n/d^{4/5}} \binom{n}{j}\left(\frac{d}{n}\right)^j Pr[I_L = 1, L \in S_i^j] \\
&= \sum_{j=1}^{\ln n/d^{4/5}} \binom{n}{j}\left(\frac{d}{n}\right) \exp\left(-d^{2/5}(j + \ln x)\right) && \text{[from (46)]} \\
&\leq \exp(-d^{4/5}\ln x) \sum_{j=1}^{\ln n/d^{4/5}} \exp\left(-d^{4/5}j + j\ln d\right) && \text{[as } \tbinom{n}{j}(d/n)^j \leq d^j\text{]} \\
&\leq \exp(-d^{4/5}\ln x - d^{4/5} + \ln d)\frac{1}{1 + e^{-d^{4/5}+\ln d}}.
\end{aligned}$$

The claim follows by noting that $\left(1 + e^{-d^{4/5}+\ln d}\right)^{-1} \leq 2$. ⋄

12.3 Proof of Lemma 14

Proof of Lemma 14: The proof is just a matter of calculations. That is

$$\begin{aligned}
S &= \sum_{j=0}^{\ln^2 n-2} jE[W^t(v_i)|\mathbb{A}, \delta(v_i) = j]Pr[\delta(v_i) = j|\mathbb{A}] \\
&\leq (1+\gamma)^{-t}\sum_{j=0}^{(1+\alpha)d-2} jPr[\delta(v_i) = j|\mathbb{A}] + d^{ct}\sum_{j=(1+\alpha)d-1}^{\ln^2 -2}(j+2)^{t+1}Pr[\delta(v_i) = j|\mathbb{A}] \\
&\leq \frac{(1+\gamma)^{-t}}{Pr[\mathbb{A}]}\sum_{j=0}^{(1+\alpha)d-2} jPr[\Delta(v_i) = j] + \frac{d^{tc}}{Pr[\mathbb{A}]}\sum_{j=(1+\alpha)d-1}^{\ln^2 -2}(j+2)^{t+1}Pr[\Delta(v_i) = j] \\
&\leq 2(1+\gamma)^{-t}\sum_{j=0}^{n} jPr[\mathcal{B}(n, d/n)] + 2d^{tc}\sum_{j=(1+\alpha)d-1}^{\ln^2 -2}(j+2)^{t+1}Pr[\Delta(v_i) = j]
\end{aligned}$$

In the last derivation we used the fact that $\Delta(v_i)$ is dominated by $\mathcal{B}(n, d/n)$ and $Pr[\mathbb{A}] \geq 1/2$. It is direct that the first summation is equal to d. As far as the second summation is regarded we use the Claim 4 we also used for the proof of Lemma 11 (See Section 11.1 for Claim 4). Thus it holds that

$$S \leq 2d(1+\gamma)^{-t} + 2d^{t(1+c)+1}\exp\left(-\alpha^2 d/5\right).$$

The lemma follows. ⋄

13 Remaining Proofs

13.1 Proof of Lemma 1

Let $\mathcal{B}_1$ be the set of blocks created from the cycles in $\mathcal{C}$ and let $\mathcal{B}_2 = \mathcal{B}\backslash\mathcal{B}_1$. It suffices to show that with probability at least $1-2n^{-3/4}$, $\mathcal{B}_1$ contains only unicyclic blocks and $\mathcal{B}_2$ contains only trees.

First we focus on the blocks in $\mathcal{B}_1$. It suffices to show that with high probability $G(n,d/n)$ is such that no two cycles in $\mathcal{C}$ are close to each other, e.g. at distance smaller than $10\frac{\ln n}{\ln^5 d}$. Then, Theorem 5 guarantees that the no two cycles will get to the same block.

If there is a pair of cycles in $\mathcal{C}$ at distance less than $10\frac{\ln n}{\ln^5 d}$, then the following should hold: There is a set of vertices S of cardinality less than $2\frac{\ln n}{\ln^2 d}$ such that the number of edges between the vertices in S is at least $|S|+1$. We are going to show that such a set does not exist in $G(n,d/n)$ with probability at least $1-n^{-3/4}$. Let D be the event that such a set exists. It holds that

$$\begin{aligned}
Pr[D] &\leq \sum_{r=1}^{2\frac{\ln n}{\ln^2 d}} \binom{n}{r}\binom{\binom{r}{2}}{r+1}\left(\frac{d}{n}\right)^{r+1} \leq \sum_{r=1}^{2\frac{\ln n}{\ln^2 d}} \left(\frac{ne}{r}\right)^r \left(\frac{r^2 e}{2(r+1)}\right)^{r+1}\left(\frac{d}{n}\right)^{r+1} && \left[\text{as } \binom{n}{r} \leq \left(\frac{ne}{r}\right)\right] \\
&\leq \frac{1}{n}\sum_{r=1}^{2\frac{\ln n}{\ln^2 d}} \left(\frac{erd}{2}\right)\left(\frac{e^2 d}{2}\right)^r \leq \frac{ed}{\ln^2 d}\cdot\frac{\ln n}{n}\sum_{r=1}^{2\frac{\ln n}{\ln^2 d}}\left(\frac{e^2 d}{2}\right)^r && [\text{as } r \leq 2\ln n/\ln^2 d] \\
&\leq \frac{ed}{\ln^2 d}\cdot\frac{\ln n}{n}\left(\frac{e^2 d}{2}\right)^{2\frac{\ln n}{\ln^2 d}} \leq n^{-3/4}
\end{aligned}$$

where the last two inequalities hold for large fixed d and large n. The above proves that part for $\mathcal{B}_1$.

So as to show that $\mathcal{B}_2$ consists of tree-like blocks we work as follows: Let some $B \in \mathcal{B}_2$ and let v be the vertex we used to created. It is direct that every path that connects v to some vertex in any of the blocks in $\mathcal{B}_1$ should contain at least one break-point (otherwise v should belong to a block in $\mathcal{B}_1$). That is, if B contains a cycle then its length should be longer than $4\frac{\ln n}{\ln^5 d}$. It suffice to show that with probability at least $1-n^{-3/4}$, every $B \in \mathcal{B}_2$ cannot contain a cycle of length $3\frac{\ln n}{\ln^5 d}$ or more.

From Theorem 5 we have that every vertex in B should be within distance at most $\frac{\ln n}{\ln^5 d}$ from v. We will show with high probability $G\left(v, \frac{\ln n}{\ln^5 d}\right)$, the induced subgraph of $G(n,d/n)$ that contains v and all the vertices within graph distance $\frac{\ln n}{\ln^5 d}$, is either a tree or unicyclic with sufficiently large probability. This implies that if there is a cycle in $G\left(v, \frac{\ln n}{\ln^5 d}\right)$ its length should not exceed $2\frac{\ln n}{\ln^5 d}+1 < 3\frac{\ln n}{\ln^5 d}$, with the same probability. Thus, we prove the part related to $\mathcal{B}_2$.

It remains to prove that with probability at least $1-n^{-3/4}$, $G\left(v, \frac{\ln n}{\ln^5 d}\right)$ is either tree or unicyclic for any v. The proof is by contradiction, e.g. assume that there is a vertex v in the graph $G(n,d/n)$ such $G\left(v, \frac{\ln n}{\ln^5 d}\right)$ contains more than one cycle. This implies that there is a set of vertices S of cardinality less than $2\frac{\ln n}{\ln^2 d}$ such that the number of edges between the vertices in S is at least $|S|+1$. We have shown above that the probability for such a set to exist in $G(n,d/n)$ is at most $n^{-3/4}$. The lemma follows.

13.2 Proof of Lemma 3

As far as (A) is regarded we use the result from [1], i.e. with probability $1-o(1)$ the chromatic number of $G(n,d/n)$ is $d/(2\ln d)$. As far as (B) is regarded we use Lemma 1 from Section 2.

As far as (C) is regarded observe the following: There is no break-point at the outer boundary of a block B that is influenced by a vertex within distance $\ln n/d^{2/5}$ inside B. Let Z_i be the number of paths L in B such that $|L| = i$ and $\prod_{u'' \in L} W(u'') > 1$. Also, let ρ_i be the probability for L to have

$\prod_{u''\in L} W(u'') > 1$. Using the tail bound in Theorem 8 (Section 6.1) we have that

$$\rho_i \leq \exp\left(-d^{4/5} i\right). \tag{47}$$

Also, let $Z = \sum_{i \geq \frac{\ln n}{d^{2/5}}} Z_i$. It is direct to see that if $Z = 0$, then the condition (C) holds. From Markov's inequality we have that

$$\begin{aligned} Pr[Z > 0] &\leq E[Z] \leq \sum_{i=\frac{\ln n}{d^{2/5}}} n^{i+1} \left(\frac{d}{n}\right)^i \rho_i \leq n \sum_{i \geq \frac{\ln n}{d^{2/5}}} d^i \rho_i \\ &\leq n \sum_{i=\frac{\ln n}{d^{2/5}}} \exp\left(-i(d^{4/5} - \ln d)\right) \qquad \text{[from 47]} \\ &\leq n^{-\frac{d^{2/5}}{2}} \sum_{i=0}^{\infty} \exp\left(-\frac{d^{4/5}}{2} i\right) \leq 2n^{-\frac{d^{2/5}}{2}}. \end{aligned}$$

That is, condition (C) is satisfied with probability at least $1 - 2n^{-\frac{d^{2/5}}{2}}$. Thus, it is direct to see that with probability $1 - o(1)$ $G(n, d/n)$ satisfies all the conditions (A), (B) and (C). The lemma follows.

13.3 Proof of Lemma 2

It is direct to show that once we have the break-points of $G(n, d/n)$ the construction of the blocks can be done in polynomial time. The lemma will follow by showing that we can distinguish whether some vertex is break point or not in polynomial time.

For a specific vertex v we need to check all the paths that start from v and are of length at most $\frac{\ln n}{d^{2/5}}$. Working as in the proof of Lemma 1 we have that with probability at least $1 - n^{-3/4}$ for every vertex v in $G(n, d/n)$ the neighbourhood we need to check is either a tree or a unicyclic graph. That is, there are at most 2 different paths between v and a vertex u at distance at most $\frac{\ln n}{d^{2/5}}$. The number of paths we need to consider is trivially upper bounded by $2n$. The lemma follows by observing that the computation of the weight of a specific path L requires $O(L)$ elementary arithmetic operations, i.e. $O(\ln n)$ operations.

13.4 Proof of Lemma 6

Assume a process where we check whether the vertices in $L = v_1, \ldots, v_{|L|}$ are right break-points, starting from v_1, then v_2, v_3 and so on. Since we are interested in right break points, so as to check v_j we don't have to examine the influence of paths getting to this vertex from its neighbor v_{j+1}.

Assume that the process is to check $v_i \in L$. Let $v_j \in L$, with $j < i$ be the last right break point we have found so far. If we haven't found any break point yet, for convention we let $j = 0$. We need to show that $\prod_{r=j+1}^{i} U(v_r)$ is bigger or equal to the influence on v_i from paths that reach the vertex either from v_{i-1} of from the neighbours of v_i outside L, i.e. the vertices in N_i. So as to show this, we use induction on the difference $i - j$.

We start from $i - j = 1$. Let $j > 0$. Then we have that the path with the maximum influence on v_j is at most 1. That is the maximum influence of paths reaching to v_i from v_j is $W(v_i)$, where $W(v_i)$ is defined in (2). Also, the maximum influence on v_i from paths that pass through its neighbours in N_i is at most $\max\{1, Q(v_i)\} \cdot W(v_i)$. Clearly among the aforementioned paths no path has influence bigger than $U(v_i)$ (i.e. $U(v_i) = \max\{1, Q(v_i)\} W(v_i)$). For $j = 0$, there are no paths reaching to v_i from v_j, i.e. we have the paths reaching from the vertices in N_i. Then, it is direct to check that the maximum influence on v_i is at most $U(v_i)$.

Let the induction hypothesis hold for $i-j=j_0$. We are going to show that it holds for $i-j=j_0+1$. By induction hypothesis we have that for v_{j+j_0} the maximum influence from a path that is reaching it either from v_{j+j_0-1} or from N_{j+j_0} is at most $\prod_{r=j+1}^{j+j_0} U(v_r) > 1$. For i such that $i-j=j_0+1$, the following holds: the influence of any path that reaches v_i through v_{i-1} (or v_{j+j_0}) cannot be larger than $W(v_i)\prod_{r=j+1}^{j+j_0} U(v_r)$ (due to induction hypothesis). Also, there is no path reaching v_i from vertices in N_i that has influence larger than $U(v_i)$. It is a matter of direct calculations to verify that none of the paths we consider has influence greater than $\prod_{r=j+1}^{j+j_0+1} U(v_r)$. The lemma follows.

Acknowledgement. The author of this work would like to thank Amin Coja-Oghlan for his comments on the work, his suggestions and the discussions.

References

[1] D. Achlioptas, and A. Naor. *The two possible values of the chromatic number of a random graph*. In proc. of 36th STOC, pp. 587-593, 2004.

[2] J. van den Berg and C. Maes. *Disagreement percolation in the study of Markov fields*. Annals of Probability **22**, pp 749-763, 1994.

[3] R. Bubley and M. Dyer. *Path Coupling: a Technique for Proving Rapid Mixing in Markov Chains.* In proc. of 38th FOCS, pp 223-231, 1997

[4] M. Dyer, A. Flaxman, A. M. Frieze and E. Vigoda. *Random colouring sparse random graphs with fewer colours than the maximum degree*. Random Struct. and Algorithms **29**, pp. 450-465, 2006.

[5] M. Dyer, A. M. Frieze, A. Hayes, E. Vigoda. *Randomly colouring constant degree graphs.* In proc. of 45th FOCS, pp 582-589, 2004.

[6] M. Dyer, A. M. Frieze and M. Jerrum. *On counting independent sets in sparse graphs*. In proc. of 40th FOCS, pp 210-217, 1999.

[7] M. Dyer, C. S. Greenhill. *On Markov chains for independent sets.* In Journal of Algorithms **35** (1), pp 17-49, 2000.

[8] C. Efthymiou. *A simple algorithm for random colouring $G(n, d/n)$ using $(2+\epsilon)d$ colours*. In proc. of SODA 2012, pp 272-280

[9] A. Frieze and E. Vigoda. *A Survey on the use of Markov Chains to Randomly Sample Colorings.* In Combinatorics, Complexity and Chance, Oxford University Press, 2007.

[10] A. M. Frieze, J. Vera. *On randomly colouring locally sparse graphs.* Discrete Mathematics & Theoretical Computer Science **8** (1), pp 121-128 2006

[11] L. A. Goldberg, R. A. Martin, M. Paterson. *Strong Spatial Mixing with Fewer Colors for Lattice Graphs.* SIAM J. Comput. **35** (2), pp 486-517 2005

[12] G. R. Grimmett and C. J. H. McDiarmid. *On colouring random graphs*. Math. Proc. Cambridge Philos. Soc., **77**, pp 313 - 324, 1975

[13] T. Hayes, J. Vera and E. Vigoda. *Randomly coloring planar graphs with fewer colors than the maximum degree.* In proc. of 39th STOC, pp 450-458, 2007.

[14] S. Janson, T. Luczak and A. Rucinski. *Random graphs*. J. Wiley and Sons, Inc. 2000.

[15] M. Jerrum. *Counting, Sampling and Integrating: Algorithms and Complexity,* Lectures in Mathematics, ETH Zurich. Birkhäuser-Verlag, 2003.

[16] M. Jerrum and A. Sinclair. *The Markov chain Monte Carlo method:an approach to approximate counting and integration.* In Approximation Algorithms for NP-hard problems, PWS 1996. (Dorit, Hochbaum ed.)

[17] F. Kelly. *Stochastic models of computer communication systems*. J. R. Statist. Soc. B **47**, (3), pp 379 - 395, 1985.

[18] F. Krzakala, A. Montanari, F. Ricci-Tersenghi, G. Semerjian, L. Zdeborova. *Gibbs states and the set of solutions of random constraint satisfaction problems.* In PNAS **104**, pp 10318-10323, 2007.

[19] D. Levin, Y. Peres and E. Wilmer. *Markov Chains and Mixing Times.* American Mathematical Society (2008).

[20] F. Martinelli, A. Sinclair and D. Weitz. Fast mixing for independent sets, colorings and other models on trees. In proc. of 15th SODA, pp 456-465, 2004.

[21] M. Molloy. *The Glauber dynamics on the colourings of a graph with large girth and maximum degree.* SIAM J. Computing **33**, pp 721-737, 2004.

[22] E. Mossel and A. Sly. *Gibbs Rapidly Samples Colorings of* $G_{n,d/n}$. In journal Probability Theory and Related Fields, **148**, (1-2), 2010.

[23] B. G. Pittel, J. Spencer and N. C .Wormald. *Sudden emergence of a giant* k*-core in a random graph.* Journal of Combinatoral Theorey B **67**, pp 111-151, 1996.

[24] E. Vigoda. *Improved bounds for sampling colorings.* Journal of Mathematical Physics, **41** (3), pp 1555-1569, 2000. A preliminary version appears in FOCS 1999.

[25] E. Vigoda. *A note on the glauber dynamics for sampling independent sets.* Elctron. J. Comb. **8** (1), pp 1-8, 2001.